Lecture Course

Algebraic Curves

Martin Schlichenmaier

Version 7.2.2017

partly put to TEX by Mathias Willibrordus Jacobus Ronk

Contents

Chapter 1

Introduction

It is not so easy to define in a mathematical but nevertheless elementary way what a curve is without using our hands. We could try:

A curve is a geometric object of dimension one.

This definition leads to additional questions: What is a *geometric object* and what is *dimension.* It will take us some time until we will be able to clarify the notions.

In this introduction we start with examples.

Example.

$$L_{a,b} := \{(x, y) \in \mathbb{R}^2 \mid y = ax - b\}, \quad a, b \in \mathbb{R}, \tag{1.1}$$

is a (straight) line in the plane $\mathbb{R}^2$, for every fixed pair of values a and b. These values we call parameters. Moreover, we could even consider instead of the field $\mathbb{R}$ any other field $\mathbb{K}$.

Example. Consider

$$Q_c := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = c\}, \quad c \in \mathbb{R}. \tag{1.2}$$

Here things get more interesting. Again c is a parameter.

1. If $c > 0$ then our Q_c will be a circle with radius $\sqrt{c}$. It is a curve in the 'usual" sense.

2. If $c = 0$ then only $\{(0, 0)\}$ is a solution of the defining equation. Hence the "naive dimension" is zero.

3. If $c < 0$ then there does not exists any solution at all, i.e. $Q_c = \emptyset$. Nevertheless the defining equation exists.

Example. We consider the defining equation above

$$x^2 + y^2 = c \tag{1.3}$$

also over the field $\mathbb{C}$ of complex numbers, meaning that we look for $(x, y) \in \mathbb{C}^2$ solving it. Recall that $\mathbb{C}$ is an algebraically closed field, and it is a field extension of $\mathbb{R}$. The two cases 2. and 3. above loose their special behavior. We denote by i the imaginary unit, i.e. $\mathrm{i}^2 = -1$.

1. $c = 0$: We write
$$x^2 + y^2 = (x + \mathrm{i}\, y)(x - \mathrm{i}\, y) = 0. \tag{1.4}$$
Hence, our D_0 (over $\mathbb{C}$) is the union of two straight (complex) lines
$$D_0 = \{(x, y) \in \mathbb{C}^2 \mid y = -\mathrm{i}\, x\} \cup \{(x, y) \in \mathbb{C}^2 \mid y = \mathrm{i}\, x\} \tag{1.5}$$
meeting each other at the unique intersection point given by the origin (which is a real solution).

2. $c < 0$: Then there again complex solutions. These are complex circles with purely imaginary radii.

Our defining equation was given over the real numbers. But that it gives curves in the usual sense could only be realized over the complex numbers, i.e. the algebraic closure of $\mathbb{R}$. We will have to take this into account.

The title of the lecture course is **Algebraic Curves**. Roughly speaking we understand by this that the defining equation is given by algebraic operations, i.e. additions, multiplications, and divisions. No transcendental functions, no exponential, no sine function, etc. are allowed in its formulation. Of course, we restrict ourselves by this, but we gain a tremendous tool box. We have algebraic techniques at our hand. Furthermore, we can consider curves over fields different from the classical fields, like $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$. Prominent examples are given by finite fields. They and the theory of algebraic curves play an important role in cryptography and coding theory.

In the next chapter we will recall the fact from algebra about field and field extensions which we need in this lecture course.

Chapter 2

Fields

In the following we recall some facts about fields. The details should have been covered in an lecture course on algebra.

We will use the symbol $\mathbb{K}$ to denote an arbitrary field. Recall that a field has two operations (1) addition "+" and (2) multiplication "$\cdot$" fulfilling the known axioms of a field. We take the convention that in our fields the multiplications are always commutative.

We denote as usual the neutral element of the addition by 0 and the neutral element of the multiplication by 1.

The *characteristic* of the field $\mathbb{K}$ (short: $\operatorname{char}\mathbb{K}$) is defined to be the smallest natural number n such that

$$\underbrace{1+1+1+\cdots+1}_{n} = 0, \tag{2.1}$$

if such a number exists. In this case the field is said to have finite characteristic. Otherwise one sets $\operatorname{char}\mathbb{K} = 0$.

Proposition 2.1. *The characteristic of a field is either zero or a prime number.*

Proof. Recall the proof as an exercise. □

Examples of fields are

1. the fields of characteristic zero: $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Q}(t)$, ...
2. the finite fields $\mathbb{F}_p$, and $\mathbb{F}_q$ with $q = p^r$, $p \in \mathbb{P}$ a prime number, $r \in \mathbb{N}$. These fields have characteristic p.

2.1 More about finite fields.

1. The field $\mathbb{F}_p$ is the residue class field $\mathbb{Z}/p\mathbb{Z}$. It has p elements. Its elements are given as $\bar{a} = a \bmod p$ with $a = 0, 1, \ldots, p-1$, with the operations

$$\bar{a} + \bar{b} := \overline{a+b}, \qquad \bar{a} \cdot \bar{b} := \overline{a \cdot b}, \tag{2.2}$$

2. The field $\mathbb{F}_q$, $q = p^r$, is the splitting field of the polynomial

$$F(X) = X^q - X \in \mathbb{F}_p[X]. \tag{2.3}$$

In fact $\mathbb{F}_q$ consists exactly of all zeros of this polynomial (and such fields exists).

Theorem 2.2. *Let $\mathbb{K}$ be a finite field, then* $\operatorname{char} \mathbb{K} = p$ *is a prime number, and there exists an $r \in \mathbb{N}$, such that $\mathbb{K} \cong \mathbb{F}_{p^r}$.*

Proof. (This is only a sketch. For details check your algebra lecture, respectively do it yourself.) As $\mathbb{K}$ has only finitely many elements its characteristic has to be finite and hence is a prime number p. Identifying the multiples of 1 in $\mathbb{K}$ with the elements of the field $\mathbb{F}_p$ we can embed $\mathbb{F}_p$ into K. Moreover, $\mathbb{K}$ is a vector space over $\mathbb{F}_p$ of certain finite dimension r. In particular $\mathbb{K}$ has exactly p^r many elements. We consider the multiplicate group $\mathbb{K}^* = \mathbb{K} \setminus \{0\}$. It has $p^r - 1$ elements. Hence by little Fermat, for every element $a \in \mathbb{K}^*$

$$(a)^{p^r-1} = 1. \tag{2.4}$$

If we multiply this with a we obtain $(a)^{p^r} = a$. Hence the elements of K are exactly the zeros of the polynomial $X^p - X$. □

2.2 Algebraic Closure

Definition 2.3. A field $\mathbb{K}$ is algebraically closed if and only if every polynomial $f \in \mathbb{K}[X]$ in one variable X, which is not a constant polynomial (i.e. $\deg f > 0$) admits a zero $\alpha \in \mathbb{K}$, i.e. $f(\alpha) = 0$.

Sometimes one calls a zero of f also a root of the polynomial f.

The field $\mathbb{R}$ is not algebraically closed, as the polynomial $X^2 + 1$ does not have a real root. In contrast the field $\mathbb{C}$ is algebraically closed.

Proposition 2.4. *A field is algebraically closed if and only if every polynomial $f \neq 0$ of $\deg f > 0$ can be written as product of linear polynomials.*

Proposition 2.5. *A finite field $\mathbb{K}$ is never algebraically closed.*

Proof. We set

$$f(X) = \prod_{\alpha \in \mathbb{K}} (X - \alpha) + 1 \in \mathbb{K}[X]. \tag{2.5}$$

This is a well-defined non-constant polynomial. Moreover for all $\beta \in \mathbb{K}$ we have $f(\beta) = 1$, as the first summand will always be zero. □

Theorem 2.6. *Let $\mathbb{K}$ be an arbitrary field. Then there exists always a field $\overline{\mathbb{K}}$ such that*

1. *$\mathbb{K} \leq \overline{\mathbb{K}}$, i.e. $\overline{\mathbb{K}}$ is a field extension of $\mathbb{K}$,*
2. *$\overline{\mathbb{K}}$ is algebraically closed,*
3. *$\overline{\mathbb{K}}$ is the smallest field with the above two properties, meaning: if $\exists \mathbb{L} : \mathbb{K} \leq \mathbb{L} \leq \overline{\mathbb{K}}$ and $\mathbb{L}$ is algebraically closed then $\mathbb{L} = \overline{\mathbb{K}}$.*

The field $\overline{\mathbb{K}}$ is called the algebraic closure of $\mathbb{K}$.

Two algebraic closures of the same field $\mathbb{K}$ will always be isomorphic.

Example. 1. The algebraic closure of $\mathbb{R}$ is $\mathbb{C}$. It can be constructed easily by adjoining to $\mathbb{R}$ a root of the irreducible polynomial $X^2 + 1$, i.e. the imaginary unit i.

2. The finite fields admit also algebraic closures. But in view of the Proposition 2.5 it will not be a finite field anymore. In particular, it cannot be obtained by adjoining only finitely many roots.

Chapter 3

Affine varieties and affine curves

3.1 The polynomial ring in n variables

Let n be a natural number. Let $X_1,\ X_2, \ldots, X_n$ be n different (formal) variables. Then we will denoted by $\mathbb{K}[X_1, X_2, \ldots X_n]$ the polynomial ring in n (commuting) variables. It can be described as vector space of finite linear combinations of monomials over the field $\mathbb{K}$.

A monomial is an expression of the type

$$X_1^{i_1} X_2^{i_2} \ldots X_n^{i_n}, \quad i_k \in \mathbb{N}_0. \tag{3.1}$$

Hence, a polynomial is a finite sum of the type

$$f(\underline{X}) = \sum a_{i_1,i_2,\ldots,i_n} X_1^{i_1} X_2^{i_2} \ldots X_n^{i_n}, \quad a_{i_1,i_2,\ldots,i_n} \in \mathbb{K}. \tag{3.2}$$

Here the sum is over the indices i_k and the coefficients $a_{i_\ldots}$ are nonzero only for finitely many combinations. For simplicity we use

$$\underline{X} = (X_1, X_2, \ldots, X_n), \quad \underline{i} = (i_1, i_2, \ldots, i_n), \quad |i| = \sum i_k, \tag{3.3}$$

and $\underline{X}^{\underline{i}}$ for the monomial (3.1).

An example of a polynomial in 4 variables is given by

$$X_1 X_2^2 + X_3^4 + 2X_1^2 X_2 X_4. \tag{3.4}$$

The vector space $\mathbb{K}[\underline{X}]$ is a commutative ring with unit 1 under the natural product of monomials and their sum.

The degree of a monomial $\underline{X}^{\underline{i}}$ is given by the sum

$$|i| := \sum_{k=1}^{n} i_k$$

of its individual degrees. The degree of a polynomial f, $\deg(f)$, is the maximal degree of non-vanishing monomials (meaning appearing monomials with non-vanishing coefficients in front of them). This is quite close to the case of polynomials in one variable with the exception that the maximal non-vanishing monomial will not be necessary unique.

Recall the following definition

Definition 3.1. Let R be a commutative ring. A non-empty subspace I is called an ideal if

$$(a) \quad I + I \subseteq I, \qquad (b) \quad R \cdot I \subseteq I. \tag{3.5}$$

Let $f_1, f_2, \ldots, f_r$ be a set of elements of R. The ideal generated by these elements is the set

$$(f_1, f_2, \ldots, f_r) \quad := \quad \{ \sum_r g_r \cdot f_r \mid g_r \in R\}. \tag{3.6}$$

It is easy to verify (exercise!) that this set is an ideal of R.

An ideal I is called finitely generated if there are finitely many $f_1, f_2, \ldots, f_r$ such that $I = (f_1, f_2, \ldots, f_r)$. A ring is called a Noetherian ring if all ideals are finitely generated [1].

From the algebra course you should know that indeed the polynomial ring in one variable is Noetherian. Indeed it is a principal domain, saying that all ideals can be generated by maximally one element.

Without proof here we quote:

Theorem 3.2. *The polynomial ring $\mathbb{K}[\underline{X}]$ in n variables over a field $\mathbb{K}$ is a Noetherian ring.*

See Section 9.2 for a proof.

Notation: If we have one, two or three variables we often also use instead of $\mathbb{K}[\underline{X}]$

$$\mathbb{K}[X], \quad \mathbb{K}[X, Y], \quad \mathbb{K}[X, Y, Z]. \tag{3.7}$$

[1] There exists quite a number of equivalent definitions for a ring being Noetherian, see Section 9.2

3.2 Affine varieties

We start with the affine space $\mathbb{A}^n(\mathbb{K})$ over the field $\mathbb{K}$. After choosing an origin O in $\mathbb{A}^n(\mathbb{K})$ and a basis in the associated vector space we can identify

$$\mathbb{A}^n(\mathbb{K}) \quad \cong \quad \mathbb{K}^n. \tag{3.8}$$

For this section we will fix such a reference frame.

With respect to our reference frame, if P is a point in the affine space it can be uniquely given by $\underline{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{K}^n$ with $\alpha_i \in \mathbb{K}$. This vector is *the vector of coordinates of the point* P. We will not always distinguish clearly between the point and its coordinate vector. As long as we stick with one reference system this will not create any problem.

Let $f \in \mathbb{K}[\underline{X}]$ be polynomial in n variables given by

$$f(\underline{X}) = \sum_{\underline{i}} a_{\underline{i}} \underline{X}^{\underline{i}}. \tag{3.9}$$

Let $P = \underline{\alpha}$ be a point. Then we can evaluate the polynomial at this point

$$f(\underline{\alpha}) = \sum_{\underline{i}} a_{\underline{i}} \underline{\alpha}^{\underline{i}} = \sum_{\underline{i}} a_{i_1, i_2, \ldots, i_n} \cdot \alpha_1^{i_1} \cdot \cdot \alpha_2^{i_2} \cdots \alpha_n^{i_n} \quad \in \mathbb{K}. \tag{3.10}$$

This defines a map (also called polynomial function) defined by f.

$$\mathbb{K} \to \mathbb{K}, \qquad \underline{\alpha} \mapsto f(\underline{\alpha}). \tag{3.11}$$

The point $\underline{\alpha}$ is called a *zero* of the polynomial f if $f(\underline{\alpha}) = 0$.

Definition 3.3. A subset A of $\mathbb{A}^n(\mathbb{K})$ is called an affine variety if there exists a subset $T \subseteq \mathbb{K}[\underline{X}]$ of polynomials such that

$$A = \mathfrak{V}(T) := \{\underline{\alpha} \in \mathbb{K}^n \mid f(\underline{\alpha}) = 0 \text{ for all } f \in T\}. \tag{3.12}$$

The set $\mathfrak{V}(T)$ is also called vanishing set of T.

Proposition 3.4. 1. *Let $S \subseteq T \subseteq \mathbb{K}[X]$ then $\mathfrak{V}(T) \subseteq \mathfrak{V}(S)$.*
2. *Let $S \subseteq \mathbb{K}[X]$ and (S) the ideal generated by S then $\mathfrak{V}(S) = \mathfrak{V}((S))$.*

Proof. Exercise. Recall that the ideal generated by S consists of all linear combination (with coefficients from $\mathbb{K}[X]$) of elements from S. □

Recall that our polynomial ring is Noetherian hence all ideals are generated by finitely many elements. By the above proposition we can use the two extremal viewpoints: affine varieties are defined by ideals or, affine varieties are defined by finitely many polynomials. Both viewpoints are equivalent.

Remark 3.5. Obviously, $\mathfrak{V}(1) = \emptyset$ and $\mathfrak{V}(0) = \mathbb{K}^n$. If one takes the set $\mathbb{K}^n$ and the affine varieties as closed sets we obtain a topology for $\mathbb{K}^n$. It is called *Zariski-Topology*. Those who know the definition of a topology are invited to show, that it is indeed a topology. (You have to show that finite unions and arbitrary intersection of closed sets are again closed.)

Example. 1. Points in $\mathbb{K}^n$ are always affine varieties. Let $\underline{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ be a point in $\mathbb{K}^n$ then

$$\mathfrak{V}(X_1 - \alpha_1, X_2 - \alpha_2, \ldots, X_n - \alpha_n) = \{\underline{\alpha}\}. \tag{3.13}$$

2. Hyperplanes are affine varieties. Hyperplanes are given as the vanishing set of one linear polynomial. For example

$$\mathfrak{V}(X_1) = \{(0, \alpha_2, \alpha_3, \ldots, \alpha_n) \mid \alpha_2, \alpha_3, \ldots, \alpha_n \in \mathbb{K}\}. \tag{3.14}$$

3. In generalization of 1. and 2. we have: Solutions of a system of linear equations are always affine varieties (known from linear algebra). Each line of the matrix equation gives a defining linear polynomial.

3.3 Values in extension fields

Above we introduced varieties by fixing a set of polynomials in $\mathbb{K}[\underline{X}]$ (indeed finitely many will do). In this sense our variety is *defined over* $\mathbb{K}$, We evaluated them by plugging in points of $\mathbb{K}^n$. We say that the variety consists of $\mathbb{K}$-valued points.

This can be extended by allowing to plug in points with coordinates from an extension field $\mathbb{L}$ of $\mathbb{K}$. Note that in this case $\mathbb{K}^n \subseteq \mathbb{L}^n$. The set of zeros are called "the $\mathbb{L}$-valued points" of the variety defined over $\mathbb{K}$. We might get more solutions. Of special importance is the case when $\mathbb{L}$ is the algebraic closure $\overline{\mathbb{K}}$ of $\mathbb{K}$.

We use the notation

$$\mathfrak{V}_{\mathbb{K}}(f_1, f_2, \ldots, f_r)[\mathbb{L}] \tag{3.15}$$

for the $\mathbb{L}$-valued points of a variety which is defined over $\mathbb{K}$ by the polynomials f_i with coefficients from $\mathbb{K}$. Now it is a subset of $\mathbb{L}^n$.

Our varieties defined in the previous sections we could also write as

$$\mathfrak{V}_{\mathbb{K}}(f_1, f_2, \ldots, f_r)[\mathbb{K}],$$

but in most cases, when the situation is clear, we will reduce this notation to the previous one.

Example. Let $f(X,Y) = X^2 + Y^2 + 1$ considered as real polynomial. We have $\mathfrak{V}_{\mathbb{R}}(X^2 + Y^2 + 1)[\mathbb{R}] = \emptyset$. If take the extension field $\mathbb{C}$ then $\mathfrak{V}_{\mathbb{R}}(X^2 + Y^2 + 1)[\mathbb{C}]$ is a complex circle in $\mathbb{C}^2$, in particular it is non-empty.

3.4 Affine Curves - Planar curves

Affine Curves are affine varieties of dimension one. Unfortunately, to develop the theory of dimension is already a quite involved task. Hence, we take here a simpler approach by restricting ourselves to the case of curves which are subset of the affine plane $\mathbb{K}^2$. The approach is restrictive as not all curves can be realized as planar curves. The most elaborated curves which we discuss here are the elliptic curves and they are planar.

Definition 3.6. An affine subset of $A^2(\mathbb{K})$ is called an affine planar curve $\mathcal{C}$ if it is defined as the vanishing set of one polynomial f of degree ≥ 1, i.e. $\mathcal{C} = \mathfrak{V}(f)$.

From the example above $\mathfrak{V}_{\mathbb{R}}(X^2 + Y^2 + 1)[\mathbb{R}] = \emptyset$. we see that it could happen that the curve (over $\mathbb{R}$) is the empty set. But if we pass to the algebraic closure $\mathbb{C} = \overline{\mathbb{R}}$ we get solutions. Being "a curve" is something which has to do with the fact that we consider the defining equation (given over $\mathbb{K}$) over the algebraic closure $\overline{\mathbb{K}}$.

Example. Let $\mathbb{K} = \mathbb{F}_3$ and consider $\mathfrak{V}(Y^2 - X^3 - X)$. In our field we have 3 elements $\{\bar{0}, \bar{1}, \bar{2}\}$. A point $(a,b) \in \mathbb{F}_3^2$ lies on the curve if and only if $b^2 = a^3 + a$. The solutions can be determined by inspecting all elements. For $a = \bar{0}$, only $b = \bar{0}$ is a solution. For $a = \bar{1}$ we have $a^3 + a = \bar{2}$. As $\bar{2}$ is not a square in $\mathbb{F}_3$, there is no solution. For $a = \bar{2}$ we have $a^3 + a = \bar{1}$. We have two solutions for b, namely $b = \bar{1}$ and $b = \bar{2}$. Hence,

$$\mathfrak{V}(Y^2 - X^3 - X) = \{(\bar{0}, \bar{0}), (\bar{2}, \bar{1}), (\bar{2}, \bar{2})\}. \tag{3.16}$$

In particular our curve has only finitely many points. This is the typical behaviour for curves defined over finite fields.

Proposition 3.7. *A curve $\mathcal{C}$ defined over a field $\mathbb{K}$ will have infinitely many points $\overline{\mathcal{C}}$ over the algebraic closure $\overline{\mathbb{K}}$. Moreover, the complimentary set $\overline{\mathbb{K}} \times \overline{\mathbb{K}} \setminus \overline{\mathcal{C}}$ will be an infinite set too.*

Proof. As this is a statement about the curve over $\overline{\mathbb{K}}$ we might assume for the proof that $\mathbb{K} = \overline{\mathbb{K}}$. The curve $\mathcal{C} = \mathfrak{V}(f)$ is defined by a non-constant polynomial $f \in \mathbb{K}[X, Y]$. By separating the variables we write

$$f = a_0 + a_1 X + \cdots + a_r X^r, \quad a_i \in \mathbb{K}[Y],\ a_r \neq 0. \tag{3.17}$$

Case 1: $(r = 0)$ then $f = a_0$ is a polynomial only in Y. As such it has only finitely many zeros. For each of this zeros we can take infinitely many values for X and obtain points on the curve. If we take a non-zero for Y we obtain in the same way infinitely many points not on the curve. Recall that an algebraically closed field always has infinitely many elements, see Proposition 2.5.
Case 2: $(r \neq 0)$ As polynomial in Y the $a_r(Y)$ has only finitely many zeros. Hence, there exists infinitely many β such that $a_r(\beta) \neq 0$. For such β we consider

$$f(X, \beta) = a_0(\beta) + a_a(\beta)X + \cdots a_r(\beta)X^r. \tag{3.18}$$

This is a non-constant polynomial in $\mathbb{K}[X]$. As $\mathbb{K}$ is algebraically closed it has zeros. If we take such a zero then $(\alpha, \beta) \in \mathcal{C}$. As we have infinitely many different β we get infinitely many points on $\mathcal{C}$. If α is not a zero we get infinitely many points in the complement. □

From the proof we can also read off that the complement is in a certain sense "bigger than the curve".

3.5 The polynomial ring is an UFD

Theorem 3.8. *The polynomial ring in n variables over a field $\mathbb{K}$ is a unique factorization domain (UFD).*

This theorem (we assume) was shown in the algebra course. Even if not, it is easy to understand what it means. The standard example of an UFD are the integers $\mathbb{Z}$. There we have the irreducible elements which are the prime numbers and the units (the invertible elements) in $\mathbb{Z}$ which is the set $\{1, -1\}$. Each integer m can be written as

$$m = \pm p_1^{i_1} \cdot p_2^{i_2} \cdots p_r^{i_r} \tag{3.19}$$

a product of a unit and powers of prime numbers. This presentation is unique up to the order of the prime factors. Strictly speaking we are also allowed to multiply the prime numbers by units.

For our polynomial ring the only invertible elements are the constant polynomials different from zero, i.e. $\mathbb{K}[\underline{X}]^* = \mathbb{K}^*$. A polynomial f is called irreducible if it cannot be written as product of two other polynomials different from the units. Warning: this notion depends on the field $\mathbb{K}$. Furthermore, to decide whether a polynomial is irreducible or not is usually not so simple. To find the decomposition (if it is reducible) is usually even harder.

We use the general statement about polynomial rings for our situation of plane affine curves.

Proposition 3.9. *Let f be a non-constant polynomial then it can be decomposed into distinct (non-equivalent) irreducible factors f_i and a unit c*

$$f = c \cdot f_1^{i_1} \cdot f_2^{i_2} \cdots f_r^{i_r}. \tag{3.20}$$

The factors f_i, $i = 1, \ldots, r$ (up to permutations of the factors and to multiplication with units) and their exponents are uniquely given.

Proposition 3.10. *Let f be a polynomial which is decomposed as (3.20), then*

$$\mathfrak{V}(f) = \mathfrak{V}(f_1 \cdot f_2 \cdots f_r). \tag{3.21}$$

Proof. If g is any polynomial and c is a unit (e.g. $c \neq 0$) then $\mathfrak{V}(g) = \mathfrak{V}(c \cdot g)$ as the constant c does not influence whether $\underline{\alpha}$ is a zero or not. Also $\mathfrak{V}(g) = \mathfrak{V}(g^2)$. Combining these two facts we get the statement. □

Definition 3.11. Given a curve $\mathcal{C} = \mathfrak{V}(f)$ with the polynomial decomposed into its different irreducible factors f_i as in (3.20), then the *degree of* $\mathcal{C}$, $\deg \mathcal{C}$, is defined to be the sum of the polynomial degrees of the different irreducible factors, i.e.

$$\deg \mathcal{C} = \sum_{i=1}^{r} \deg f_i. \tag{3.22}$$

For any curve $\mathcal{C} = \mathfrak{V}(f)$ we can ignore multiple factors in the decomposition of f. In such a way we obtain another polynomial f' such that $\mathcal{C} = \mathfrak{V}(f')$ which is minimal with respect to the degree. This polynomial we call *defining polynomial* or *minimal polynomial* for $\mathcal{C}$.

Remark 3.12. In more advanced algebraic geometry one likes to talk about geometric objects with multiplicities. In this more general setup $\mathfrak{V}(f)$ and $V(f^2)$ (now considered as geometric schemes) would make a difference.

Proposition 3.13. *Let f_1 and f_2 be two polynomials and $f = f_1 \cdot f_2$ then*

$$\mathfrak{V}(f) = \mathfrak{V}(f_1) \cup V(f_2). \tag{3.23}$$

Proof. Let $(a, b) \in \mathfrak{V}(f)$ then

$$0 = f(a, b) = f_1(a, b) \cdot f_2(a, b) \implies f_1(a, b) = 0 \text{ or } f_2(a, b) = 0. \tag{3.24}$$

Hence $(a, b) \in \mathfrak{V}(f_1)$ or $(a, b) \in \mathfrak{V}(f_2)$. The opposite is clear. □

This has the following consequence

Proposition 3.14. *Let f be a polynomial which is decomposed as (3.20), then*

$$\mathfrak{V}(f) = \mathfrak{V}(f_1) \cup \mathfrak{V}(f_2) \cup \cdots \cup V(f_r). \tag{3.25}$$

For example, the affine variety $\mathfrak{V}(X \cdot Y)$ decomposes as $\mathfrak{V}(X) \cup \mathfrak{V}(Y)$. It is the union of the coordinate axes.

Definition 3.15. Let $\mathcal{C}$ be a a (planar) affine curve. It is called *reducible* if and only if there exists two curves $\mathcal{C}_1$ and $\mathcal{C}_2$, $\mathcal{C}_1 \neq \mathcal{C}_2$ such that $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$.

Otherwise $\mathcal{C}$ is called *irreducible*.

Next we want to examine closer the relation between reducibility of curves and defining polynomials. This correspondence in both ways will only work if we consider curves and polynomials over algebraically closed fields.

Example. We consider $\mathfrak{V}((X^2 + Y^2)(X - Y))$ over $\mathbb{R}$ (and as variety over $\mathbb{R}$). The polynomial there does not have multiple components. Hence it is the defining polynomial. Clearly it is not irreducible. But $\mathfrak{V}((X^2 + Y^2)(X - Y)) = \mathfrak{V}(X - Y)$ and this is a straight line which is irreducible. If we consider the variety over $\mathbb{C}$, meaning that we allow $\mathbb{C}$-valued points, then the variety decomposes into the union of three lines, hence it is also reducible.

Definition 3.16. (a) A polynomial $f \in \mathbb{K}[X, Y]$ is called absolutely irreducible if and only if f is irreducible as polynomial in $\overline{\mathbb{K}}[X, Y]$.
(b) A curve $\mathcal{C} = \mathfrak{V}_{\mathbb{K}}(f)[\mathbb{K}]$ is called absolutely irreducible if and only if the curve $\mathfrak{V}_{\mathbb{K}}(f)[\overline{\mathbb{K}}]$ is irreducible.

Proposition 3.17. *A curve $\mathcal{C}$ is absolutely irreducible if and only if the defining polynomial f for $\mathcal{C}$ is absolutely irreducible.*

Before we say something on the proof, we formulate the following lemma which we will discuss further down.

Lemma 3.18. *Let $\mathbb{K}$ be algebraically closed. Let $f, g \in \mathbb{K}[X, Y]$ without multiple factors. Then*

$$\mathfrak{V}(f) = \mathfrak{V}(g) \quad \Longrightarrow \quad \exists c \in \mathbb{K}^* : g = c \cdot f. \tag{3.26}$$

This says that over an algebraically closed field the defining polynomial of a curve is uniquely given (up to multiplications with units).

Proof. (of Proposition 3.17) Without restriction we may assume that we already work over an algebraically closed field. Let $\mathcal{C}$ be irreducible with defining polynomial f. Recall that by convention it will not have multiple factors. Assume that $f = f_1 \cdot f_2$ then $\mathfrak{V}(f) = \mathfrak{V}(f_1) \cup \mathfrak{V}(f_2)$. Hence $\mathcal{C}$ is the union of two curves $\mathcal{C}_1$ and $\mathcal{C}_2$. As $\mathcal{C}$ is irreducible it has to be one of them, say $\mathcal{C}_1$. Then $\mathfrak{V}(f) = \mathfrak{V}(f_1)$ this contradicts Lemma 3.18. Vice versa, let $\mathcal{C} = V(f)$ with f irreducible. Assume that $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$ is a non-trivial decomposition into curves. Let $\mathcal{C}_i = \mathfrak{V}(f_i)$, and note that f_1 and f_2 are not equivalent. For the union we get $\mathcal{C} = \mathfrak{V}(f_1 \cdot f_2) = \mathfrak{V}(f)$. Lemma 3.18 again says that $f = c \cdot f_1 \cdot f_2$ which is a contradiction to the fact that f is irreducible. □

If we give a curve $\mathcal{C} = \mathfrak{V}(f)$ by its defining polynomial f we can use the decomposition of f into its irreducible factors f_i and get (over an algebraically closed field) a decomposition of $\mathcal{C}$ into its irreducible components:

Theorem 3.19. *(over an algebraically closed field) Every algebraic curve $\mathcal{C}$ can be decomposed into*

$$\mathcal{C} = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_r \tag{3.27}$$

with C_i irreducible and distinct curves. The decomposition corresponds to the decomposition of the defining polynomial f into irreducible factors $\mathcal{C}_i = \mathfrak{V}(f_i)$. It is unique up to reordering.

3.5.1 The Lemma 3.18

For this subsection let $\mathbb{K}$ be an algebraically closed field. Given a subset $A \subseteq \mathbb{A}^n(\mathbb{K})$ we assign to it

$$\mathcal{I}(A) := \{ f \in \mathbb{K}[\underline{X}] \mid \forall \underline{\alpha} \in A : f(\underline{\alpha}) = 0 \}. \tag{3.28}$$

This set is called the vanishing ideal of $\mathbb{A}$.

Exercise: $\mathcal{I}(A)$ is an ideal in $\mathbb{K}[\underline{X}]$.

The question is how are $\mathfrak{V}$ and $\mathcal{I}$ related.

Proposition 3.20. *Let V be a variety. Then*

$$\mathfrak{V}(\mathcal{I}(V)) = V. \tag{3.29}$$

Proof. Obviously, $V \subseteq \mathfrak{V}(I(V))$ as the elements of $\mathcal{I}(V)$ on V vanish. But if $J = (f_1, f_2, \ldots, f_k)$ is a defining ideal for V (i.e. $V = \mathfrak{V}(J)$) then $J \subseteq \mathcal{I}(V)$. This implies that $V = \mathfrak{V}(J) \supseteq \mathfrak{V}(\mathcal{I}(V))$. Hence equality. □

Proposition 3.21. *Let I be an ideal in $\mathbb{K}[\underline{X}]$ then*

$$\mathcal{I}(\mathfrak{V}(I)) \supseteq I. \tag{3.30}$$

Proof. From the definition of $\mathfrak{V}(I)$ it follows that for all $f \in I$, $f(\underline{\alpha}) = 0$, hence $f \in \mathcal{I}(\mathfrak{V}(I))$. □

Given an ideal I we define its *radical* $Rad(I)$ to be

$$Rad(I) := \{ f \in \mathbb{K}[\underline{X}] \mid \exists r \in \mathbb{N} : f^r \in I \}. \tag{3.31}$$

For example for $n = 1$ and $I = (X^2)$ we get $Rad(I) = (X)$.

Without proof (see [KK]) we quote

Theorem 3.22. *(Hilbertscher Nullstellensatz (HNS)). Over an algebraically closed field $\mathbb{K}$*

$$\mathcal{I}(\mathfrak{V}(I)) = Rad(I). \tag{3.32}$$

If $I = Rad(I)$ then the ideal I is called radical ideal. A more refined analysis shows that the set of radical ideals of $\mathbb{K}[\underline{X}]$ is in 1:1 correspondence to varieties in $\mathbb{A}^n(\mathbb{K})$ (for algebraically closed fields $\mathbb{K}$). The bijective (inverse) maps are given by $\mathcal{I}$ and by $\mathfrak{V}$. For example $\mathcal{I}(\mathfrak{V}(Rad(I)) = Rad(I)$, the vanishing ideal is always a radical ideal, etc.

Proof. (of Lemma 3.18) Let $V(f) = V(g)$. Then by the HNS $g \in \mathcal{I}(f)) = Rad((f))$. This says there exists a $r \in \mathbb{N}$ such that $g^r = k \cdot f$, with $k \in \mathbb{K}[\underline{X}]$. As f does not have multiple components $g = k \cdot f$. The same chain of arguments for f gives $f = l \cdot g$. This implies $k \cdot l = 1$ and hence they are units. This was exactly to be shown. □

Remark 3.23. 1. Lemma 3.18 (and its proof above) is valid for all dimensions.
2. In dimension two the use of HNS can be circumvented, see e.g. [Kunz].
3. In dimension two all nontrivial ideals are given as $I = (f)$ with a polynomial f. Then I is a radical ideal if and only if f does not have multiple irreducible factors.

3.5.2 Intersections

Above we saw that finite union of curves are again curves. One just multiplies the defining polynomials. The situation is more complicated for the intersection of two (plane) curves. In case that both curves are irreducible and different the intersection will not be a curve anymore. It might be empty or consist of a collection of points. Indeed the intersection is always finite. Without proof we quote here:

Proposition 3.24. *Let f, g be non-constant relatively prime polynomials from $\mathbb{K}[X, Y]$ then $V(f) \cap V(g)$ is a finite set (possibly empty).*

3.6 Singularities

To "linearize' a curve $\mathcal{C}$ at a point P one uses tangents at the point along the curve. From elementary analysis and geometry you should remember that tangents are related to derivatives of the defining equations at the point. But there might exist "bad points" of the curve where it is not possible to define the tangent.

These points will be the singular points of the curve. In this section we will only give the definition. The relation to tangent lines will be taken up later.

Tangents are defined with the help of derivatives of the defining equations, i.e. by differential quotients. In analysis they are defined as limits of difference equations. For arbitrary fields we do not have this basic tool of analysis at hand. Nevertheless we know the rules

$$(X^n)' = n \cdot X^{(n-1)}, \quad \text{etc.} \tag{3.33}$$

These rules are purely algebraic rules and make sense (as a definition) for our polynomials.

Given a monomial $\underline{X}^{\underline{i}} = X_1^{i_1} \cdots X_n^{i_n}$ we define the *formal partial derivative* with respect to the variable X_j to be

$$\frac{\partial \underline{X}^{\underline{i}}}{\partial X_j} := \begin{cases} 0, & i_j = 0 \\ i_j X_1^{i_1} \cdots X_j^{i_j - 1} \cdots X_n^{i_n}, & i_j \geq 1. \end{cases} \tag{3.34}$$

Furthermore, we extend this $\mathbb{K}-$linear to all of $\mathbb{K}[\underline{X}]$. For sure $\frac{\partial f}{\partial X_j}$ is again a polynomial and we have

Proposition 3.25. *If* $\frac{\partial f}{\partial X_j} \neq 0$ *then* $\deg \frac{\partial f}{\partial X_j} = \deg f - 1$. *Furthermore*

$$\frac{\partial (f \cdot g)}{\partial X_j} = \frac{\partial f}{\partial X_j} \cdot g + f \cdot \frac{\partial g}{\partial X_j}. \tag{3.35}$$

The relation (3.35) is called Leibniz rule.

Proof. This is an exercise. Clearly, it is enough to do it for monomials. □

Remark 3.26. A word of warning is in order. Consider the monomial X^2 in $\mathbb{F}_2[X]$. Then

$$\frac{\partial X^2}{\partial X} = 2 \cdot X = 0 \tag{3.36}$$

as $\bar{2} = \bar{0}$. This shows that the rule that if all partial derivatives are vanishing then f is a constant is only true in characteristic zero.

Definition 3.27. (a) Let $\mathcal{C} = V_{\mathbb{K}}(f)[\mathbb{K}]$ be an affine planar curve defined over the field $\mathbb{K}$ with values in $\mathbb{K}$. A point $(a, b) \in \mathbb{K}^2$ is called a *singular point* of $\mathcal{C}$ if and only if

$$f(a, b) = 0, \qquad \frac{\partial f}{\partial X}(a, b) = 0, \qquad \frac{\partial f}{\partial Y}(a, b) = 0. \tag{3.37}$$

(b) A curve $\mathcal{C} = V_{\mathbb{K}}(f)[\mathbb{K}]$ is called *non-singular* if and only if $V_{\mathbb{K}}(f)[\overline{\mathbb{K}}]$ does not have any singular points $(a, b) \in \overline{\mathbb{K}}$.

To decide whether a curve is non-singular we have to consider the curve over the algebraic closure of the defining field.

The equations (3.37) are algebraic equations with coefficients from $\mathbb{K}$, hence the following is obvious.

Proposition 3.28. *Let* $(a, b) \in \mathbb{K}^2$, *and let* $\mathbb{L}$ *be an extension field of* $\mathbb{K}$. *Then* (a, b) *is a singular point of* $\mathcal{C} = \mathfrak{V}_{\mathbb{K}}(f)[\mathbb{K}]$ *if and only if* (a, b) *considered as point in* $\mathbb{L}^2$ *is a singular point of* $\overline{\mathcal{C}} = \mathfrak{V}_{\mathbb{K}}(f)[\mathbb{L}]$.

Example.

$$f(X, Y) = Y^2 - X^4 - 2X^2 - 1. \tag{3.38}$$

Consider this curve over $\mathbb{R}$. We calculate

$$\frac{\partial f}{\partial Y} = 2Y, \quad \frac{\partial f}{\partial X} = -4X(X^2 + 1). \tag{3.39}$$

Assume that $(x, y) \in \mathbb{R}^2$ is a singular point, then from these equations it follows that $(x, y) = 0$. But $f(0, 0) = -1 \neq 0$. Hence $\mathfrak{V}_{\mathbb{R}}[\mathbb{R}]$ does not have singular points. Next we have to consider the curve over $\mathbb{C}$, the algebraic closure of $\mathbb{R}$. Still $y = 0$ is the only possible solution for y. But now we have also to consider $(\mathrm{i}\,, 0)$ and $(-\mathrm{i}\,, 0)$. Indeed $f(\pm\mathrm{i}\,, 0) = 0$. Hence $\mathfrak{V}_{\mathbb{R}}[\mathbb{C}]$ has these two points as singular points and the curve is singular.

Example.

$$f(X, Y) = Y^2 - X^3 + X = Y^2 - X(X + 1)(X - 1). \tag{3.40}$$

Over $\mathbb{R}$ we calculate

$$\frac{\partial f}{\partial Y} = 2Y, \quad \frac{\partial f}{\partial X} = -3X^2 + 1. \tag{3.41}$$

From the first equation $\frac{\partial f}{\partial Y}(x, y) = 0$ we conclude that for a singular point (x, y), $y = 0$ is necessary. But then $f(x, y) = 0$ implies that $x = 0, 1$ or -1. None of these values satisfies $\frac{\partial f}{\partial X}(x, 0) = 0$. Hence, there does not exist any singular points. This is also true for the algebraic closure $\mathbb{C}$. In particular this curve is non-singular.

Example.

$$f(X, Y) = Y^2 - X^3. \tag{3.42}$$

Now

$$\frac{\partial f}{\partial Y} = 2Y, \quad \frac{\partial f}{\partial X} = -3X^2. \tag{3.43}$$

Obviously the point $(0, 0)$ is a singular point (the only one). This curve is called cuspidal cubic, see Figure 3.1

Example.

$$f(X, Y) = Y^2 - X^3 - X^2 = Y^2 - X^2(X + 1). \tag{3.44}$$

Now

$$\frac{\partial f}{\partial Y} = 2Y, \quad \frac{\partial f}{\partial X} = -3X^2 - 2X = -X(3X + 2). \tag{3.45}$$

The point $(0, 0)$ is a singular point (the only one). This curve is called nodal cubic, see Figure 3.2

Proposition 3.29. *Let $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$ be a reducible curve. Assume that there is a point $P \in \mathcal{C}_1 \cap \mathcal{C}_2$ in the intersection, then P is a singular point of $\mathcal{C}$.*

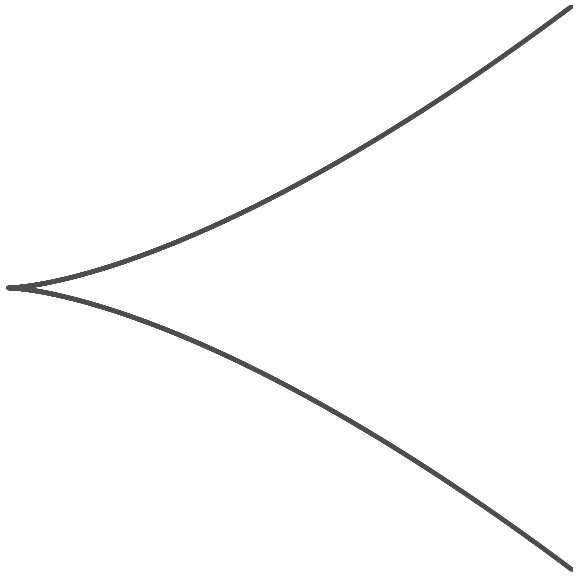

Figure 3.1: cuspidal cubic.

Proof. Let $\mathcal{C} = \mathfrak{V}(f) = \mathfrak{V}(f_1 \cdot f_2)$ be the decomposition. Let P given by the coordinates (a, b). As P is on $\mathcal{C}$, $\mathcal{C}_1$ and $\mathcal{C}_2$ we have $f(a, b) = f_1(a, b) = f_2(a, b) = 0$. For the derivative we get

$$\frac{\partial f}{\partial X} = \frac{\partial f_1}{\partial X} \cdot f_2 + \frac{\partial f_2}{\partial X} \cdot f_1. \tag{3.46}$$

If we plug in (a, b) this gives 0. The same is true for the derivative with respect to Y. Hence (a, b) is a singular point. □

Note the existence of a point of intersection is crucial for the proof.

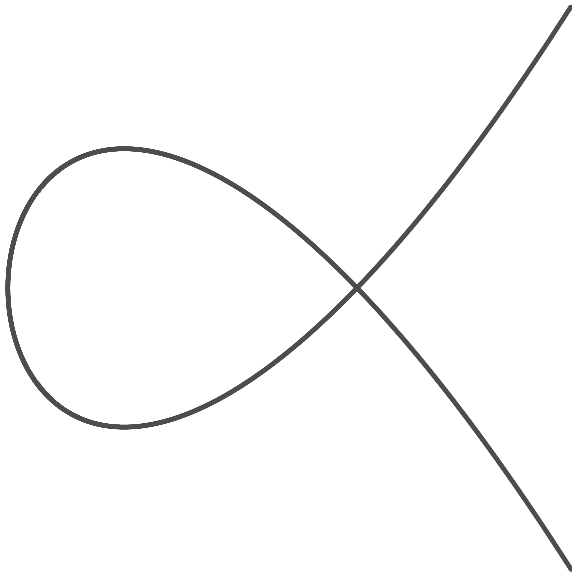

Figure 3.2: nodal cubic.

Chapter 4

Projective Varieties

Affine varieties have certain disadvantages:

1. Affine varieties over the field $\mathbb{C}$ are topological sets, more precisely they carry the usual topology as subsets of $\mathbb{R}^{2n} = \mathbb{C}^n$. But they are not compact in this topology. In analysis if one deals with compact sets one has much stronger results. For example, a continuous function defined on a compact set has a maximum and a minimum. Many more things are valid.

2. In affine geometry one has too many case distinction . For example, given two different straight lines. They will not always meet. They only meet if they are not parallel.

We will introduce now the projective space, projective varieties and projective curves. In projective geometry we will have the result that two straight lines will always have a point of intersection. Moreover over an algebraically closed field two projective curves in the plane will always have a non-empty intersection.

4.1 Projective Space

We consider the vector space $\mathbb{K}^{n+1}$ and define the following relation on pairs of points: $\alpha = (\alpha_0, .., \alpha_n)$ and $\beta = (\beta_0, .., \beta_n)$

$$\alpha \sim \beta \quad \text{if and only if} \quad \exists \lambda \in \mathbb{K}^* : \beta = \lambda \cdot \alpha \ .$$

Exercise: Show this is an equivalence relation. Recall what one has to show reflexivity, symmetry and transitivity.

The equivalence class of $\alpha \in \mathbb{K}^{n+1}$ is defined as

$$[\alpha] := \{\beta \in \mathbb{K}^{n+1} \mid \beta \sim \alpha\}\,.$$

In can also be written as

$$[\alpha] = \mathbb{K}^* \cdot \alpha\,.$$

The zero vector $0 \in \mathbb{K}^{n+1}$ is special in the sense that

$$[0] = \{0\}\,.$$

Definition 4.1. The projective space over $\mathbb{K}$ of dimension n is defined to be the set of equivalence classes of vectors from $\mathbb{K}^{n+1}\backslash\{0\}$.
In symbols

$$\mathbb{P}^n(\mathbb{K}) = (\mathbb{K}^{n+1}\backslash\{0\})/\sim\ .$$

To denote a point $[\alpha]$ in projective space $\mathbb{P}^n(K)$ we use the $\alpha = (\alpha_0, .., \alpha_n)$. as homogeneous coordinates $[\alpha] = (\alpha_0 : .. : \alpha_n)$.. Note that the homogeneous coordinates are not uniquely defined as there exists $\beta \neq \alpha$ with $[\beta] = [\alpha]$. Hence

$$[\alpha] = (\beta_0 : ... : \beta_n)$$

is also a set of homogeneous coordinates for the same point. The relation between these two coordinates are that there exists a $\lambda \in \mathbb{K}^*$ such that

$$\beta_i = \lambda \cdot \alpha_i \qquad \forall i \in \{0, 1, ..., n\}\,.$$

To give an example, over a field $\mathbb{K}$ of $\mathrm{char}(\mathbb{K}) \neq 2$

$$(1 : 1 : ... : 1) = (2 : 2 : ... : 2)\,.$$

Question: Why do we have to exclude $\mathrm{char}(\mathbb{K}) \neq 2$?

We introduce the following subsets $U_i \subseteq \mathbb{P}^n(\mathbb{K})$, $i = 0, 1, ..., n$

$$U_i := \{\alpha = (\alpha_0 : .. : \alpha_n) \in \mathbb{P}^n \mid \alpha_i \neq 0\}\,.$$

The subset U_i can be identified with the n-dimensional affine space $\mathbb{K}^n$ by normalizing in the homogeneous coordinates the value $\alpha_i = 1$ for the points in U_i.

As an example we consider $i = 0$

$$\mathbb{A}^n(\mathbb{K}) \cong \mathbb{K}^n \longrightarrow U_0 \subseteq \mathbb{P}^n(\mathbb{K})$$

$$(\alpha_1, .., \alpha_n) \longmapsto (1 : \alpha_1 : .. : \alpha_n) .$$

This is a bijection.

What are the points in the complement of U_0 ?
These are the points with homogeneous coordinate $(0 : \alpha_0 : .. : \alpha_n)$. Of course the condition remains that this vector should be different from 0. Hence, if we strip off the first component we get exactly the projective space of one dimension less, i.e. $\mathbb{P}^{n-1}(\mathbb{K})$.

Hence

$$\mathbb{P}^n(\mathbb{K}) = U_0 \cup \mathbb{P}^{n-1}(\mathbb{K}) = \mathbb{A}^n(\mathbb{K}) \cup \mathbb{P}^{n-1}(\mathbb{K})$$

where the union is a disjoint union. What has been done for $i = 0$ could be done with respect to every $i \in \{0, 1, ..., n\}$

$$\mathbb{P}^n(\mathbb{K}) = U_i \cup \mathbb{P}^{n-1}(\mathbb{K}) .$$

The points of $\mathbb{P}^n(\mathbb{K})$ which lie in $\mathbb{P}^{n-1}(\mathbb{K})$ with respect to this decomposition are called points at infinity (with respect to U_i), the points in U_i are called the affine points. Globally we have

$$\mathbb{P}^n(\mathbb{K}) = \bigcup_{i=0}^{n} U_i, \quad U_i \cong \mathbb{A}^n(\mathbb{K}) .$$

This is clear as by definition for every $[\alpha] \in \mathbb{P}^n(\mathbb{K})$ there exists an i such that $\alpha_i \neq 0$, hence $[\alpha] \in U_i$.

Example. $\mathbb{P}^1(\mathbb{K})$ the projective line.
We have by definition

$$\mathbb{P}^1(\mathbb{K}) = \{(\alpha_0 : \alpha_1) \mid \alpha_0, \alpha_1 \in \mathbb{K}, (\alpha_0, \alpha_1) \neq (0, 0)\} .$$

Our affine points are

$$U_0 := \{(1 : \alpha_1) \mid \alpha_1 \in \mathbb{K}\} \cong \mathbb{K} ,$$

$$U_1 := \{(\alpha_0 : 1) \mid \alpha_0 \in \mathbb{K}\} \cong \mathbb{K} .$$

For their intersection we obtain

$$U_0 \cap U_1 = \{(1 : \alpha_1) \mid \alpha_1 \in \mathbb{K}, \alpha_1 \neq 0\} \cong \mathbb{K} \backslash \{0\} .$$

For their points at infinity

$$\mathbb{P}^1 \backslash U_0 = \{(0 : \alpha_1)\} = \{(0 : 1)\}\,.$$

Hence it is exactly one point $(0 : 1)$ this is the point with respect to U_0. The point with respect to U_1 is $(1 : 0)$.

Remark 4.2. In the case $\mathbb{K} = \mathbb{C}$, the projective line $\mathbb{P}^1(\mathbb{C})$ can be identified with the two-sphere where the point at ∞ is the north pole.

Example. The projective plane $\mathbb{P}^2(\mathbb{K})$.

Let us consider $i = 2$. Then

$$U_2 := \{(\alpha_0 : \alpha_1 : \alpha_2) \mid \alpha_0, \alpha_1 \in \mathbb{K}, \alpha_2 \neq 0\} = \{(\alpha_0 : \alpha_1 : 1) \mid \alpha_0, \alpha_1 \in \mathbb{K}\} \cong \mathbb{A}^2(\mathbb{K})\,.$$

The complement is the "line at ∞"

$$\mathbb{P}^2 \backslash \mathbb{A}^2 = \{(\alpha_0, \alpha_1, 0) \mid \alpha_0 \neq 0\,, \text{or } a_1 \neq 0\} \cong \mathbb{P}^1\,.$$

4.2 Projective Varieties

Affine varieties were defined as zero sets of polynomials. This cannot be extended without modifications to the projective situation. Let us start with a polynomial f in $(n+1)$ variables and a point $[\alpha] \in \mathbb{P}^n$ with homogeneous coordinates
$\alpha = (\alpha_0 : ... : \alpha_n)$.

If we fix homogeneous coordinates for $[\alpha]$ we can plug-in α in $f(x)$ and get $f(\alpha) \in \mathbb{K}$. If β are another homogeneous coordinates for the same point $[\alpha]$ we will (in general) obtain different values $f(\beta)$. Hence arbitrary polynomials are not functions for the point in projective space.

Let $M = X_0^{i_0} \cdot X_1^{i_1} \cdot \cdots \cdot X_n^{i_n}$ be a monomial of degree $d = \sum_{k=0}^{n} i_k$. Recall that the set of different monomials constitute a basis of $\mathbb{K}[X_0, ..., X_n]$.

Definition 4.3. A polynomial $f \in \mathbb{K}[X_0, ..., X_{n+1}]$ is called *homogeneous* of degree d if and only if f is a sum of monomials of the same degree d.

An example is given by

$$f(X, Y, Z) = X^2Y + Y^3 + ZXY$$

which is a homogeneous polynomial of degree 3.

Now we plug in α and β for $[\alpha] = [\beta]$ into such a monomial M of degree d. Recall that $[\alpha] = [\beta]$ means that $\beta = \lambda \cdot \alpha$,with $\lambda \in \mathbb{K}^*$

$$M(\beta) = (\lambda \cdot \alpha_1)^{i_0} \cdot (\lambda \cdot \alpha_2)^{i_1} \cdot \cdots \cdot (\lambda \cdot \alpha_n)^{i_n} = \lambda^{\sum_{k=0}^{n} i_k} \cdot M(\alpha) = \lambda^d M(\alpha)\,.$$

If f is a homogeneous polynomial of degree d then

$$f(\beta) = \lambda^d f(\alpha)\,.$$

This is still not a function but now at least we have

$$f(\alpha) = 0 \Longleftrightarrow f(\beta) = 0\,.$$

This means that the zeros of f are well-defined. For an arbitrary polynomial f even this is not the case.

Definition 4.4. A subset V of $\mathbb{P}^n(\mathbb{K})$ is called projective variety if and only if there exist k homogeneous polynomials f_i (not necessarily of the same degree d_i) such that

$$\begin{aligned} V &= \mathfrak{V}(f_1, ..., f_k) \\ &= \mathfrak{V}_{\mathbb{K}}(f_1, ..., f_k)[\mathbb{K}] \\ &= \{[\alpha] \in \mathbb{P}^n(\mathbb{K}) \mid f_1(\alpha) = f_2(\alpha) = ... = f_k(\alpha) = 0\}\,. \end{aligned}$$

A special case is

Definition 4.5. A subset C of the projective plane $\mathbb{P}^2(\mathbb{K})$ is called (planar) projective curve if and only if there exist a homogeneous polynomial $f \in \mathbb{K}[X, Y, Z]$ of degree $d > 0$ such that

$$C = \mathfrak{V}(f)\,.$$

If f does not have multiples factors then C is called a curve of degree d.

Next we want to examine the relation between affine and projective curves.

Let $\tilde{C}$ be a projective curve defined by the polynomial $\tilde{f} \in \mathbb{K}[X, Y, Z]$,
i.e. $\tilde{C} = \mathfrak{V}(\tilde{f})$. Let us take as affine subset

$$U_2 = \{(\alpha_0 : \alpha_1 : 1) \mid \alpha_0, \alpha_1 \in \mathbb{K}\}\,.$$

and consider

$$\tilde{C} \cap U_2 = \{(\alpha_0 : \alpha_1 : 1) \mid \tilde{f}(\alpha_0, \alpha_1, 1) = 0\}\,.$$

If we identify U_2 with $\mathbb{K}^2$ and set

$$f(X, Y) := \tilde{f}(X, Y, 1) \in \mathbb{K}[X, Y]\,.$$

We obtain a polynomial in 2 variables (witch in general will not be homogeneous anymore). Then

$$\tilde{C} \cap U_2 \cong C = \mathfrak{V}(f)\,,$$

which is an affine curve in $\mathbb{K}^2$. Consequence: By restricting a projective curve to an affine part of the projective plane we obtain an affine curve.

Now we go in the opposite direction. We start with an affine curve $C = \mathfrak{V}(f)$ with $f \in \mathbb{K}[X, Y]$. Let d be the maximal degree of monomials appearing in f. Let M be one of the monomials in f of degree $d(M)$, then we augment M by multiplying it with $Z^{d-d(M)}$.

In total we obtain then a polynomial $\tilde{f}(X, Y, Z)$ witch is homogeneous of degree d.

Example. Let

$$f(X, Y) = Y^2 - 4X^3 - X$$

this is a polynomial of degree 3. Hence the maximal degree of monomials is 3. We obtain

$$\tilde{f}(X, Y, Z) = Y^2 Z - 4X^3 - XZ^2\,.$$

By construction $\tilde{f}(X, Y, 1) = f(X, Y)$.

Let $\tilde{C}$ be the projective curve defined by $\tilde{f}$ i.e. $\tilde{C} = \mathfrak{V}(\tilde{f})$. By construction

$$\tilde{C} \cap U_2 = \mathfrak{V}(\tilde{f}) \cap U_2 \cong C = \mathfrak{V}(f)\,.$$

The curve $\tilde{C}$ obtained in this way is called *projective completion* of C.

Example. Let $f_1(X, Y) = Y - a$ with $a \neq 0$, $f_2(X, Y) = Y$ and $L_1 = \mathfrak{V}(f_1), L_2 = \mathfrak{V}(f_2)$, the two corresponding affine curves. In fact these are 2 parallel lines (distinct). Now $\tilde{L_1} \cap \tilde{L_2} = \emptyset$. We pass from $\mathbb{K}^2$ to $\mathbb{P}^2(\mathbb{K})$.
For this we have to homogenize f_1, f_2 and get

$$\tilde{f}_1(X, Y, Z) = Y - aZ, \qquad \tilde{f}_2(X, Y, Z) = Y$$

Let $\tilde{L_1}$ and $\tilde{L_2}$ be the projectively completed curves.
We calculate

$$\begin{aligned}\tilde{L_1} \cap \tilde{L_2} &= \{(\alpha : \beta : \gamma) \mid \tilde{f}_1(\alpha, \beta, \gamma) = 0 = \tilde{f}_2(\alpha, \beta, \gamma)\} \\ &= \{(\alpha : \beta : \gamma) \mid \beta - a \cdot \gamma = \beta = 0\} \\ &= \{(1 : 0 : 0)\}\,.\end{aligned}$$

Hence we obtain a point of intersection. Of course this point does not lie on the affine part but on the line at infinity.

Exercise: Do the same for the affine lines which still are parallel but now have a slope of $r \in \mathbb{K}$, i.e.

$$f_1(X, Y) = Y - rX - a \quad f_2(X, Y) = Y - rX$$

As point of intersection of the projectively completed lines you will calculate the point $(1 : r : 0)$. The intersection point on the line at infinity will be given by the slope of the affine lines.

4.3 Singularities

Definition 4.6. Let $f \in \mathbb{K}[X, Y, Z]$, be homogeneous of degree d.

1. (a) $C = \mathfrak{V}(f)[\mathbb{K}]$ a curve of degree d over a field $\mathbb{K}$ for which $\operatorname{char} \mathbb{K} \nmid d$ is called *singular* at the point $P = [(\alpha : \beta : \gamma)]$ if and only if all partial derivatives are zero at the point P, i.e.

$$\frac{\partial f}{\partial X}(P) = \frac{\partial f}{\partial Y}(P) = \frac{\partial f}{\partial Z}(P) = 0\,. \tag{4.1}$$

   The point P is called singular point of $\mathfrak{V}(f)$.

   (b) In case that $\operatorname{char} \mathbb{K}$ divides d we put $f(P) = 0$ as additional condition for P being a singular point of C..

2. The curve $\mathfrak{V}(f)[\mathbb{K}]$ is called *non-singular* if and only if $\mathfrak{V}(f)[\overline{\mathbb{K}}]$ does not have any singular points.

There are some remarks necessary:

Proposition 4.7. *1. If f is homogeneous of degree d then $\frac{\partial f}{\partial X}, \frac{\partial f}{\partial Y}, \frac{\partial f}{\partial Z}$ are homogeneous polynomials of degree $d-1$ (or zero).*

2. (Relation of Euler) If f is of degree d then

$$d \cdot f = X \cdot \frac{\partial f}{\partial X} + Y \cdot \frac{\partial f}{\partial Y} + Z \cdot \frac{\partial f}{\partial Z}\,. \tag{4.2}$$

Proof. **1.** This we showed already before, see Proposition 3.25.
2. Let m be a monomial of degree d , i.e.

$$m = X^{d_1} Y^{d_2} Z^{d_3} \quad d = d_1 + d_2 + d_3\,,$$

then

$$\frac{\partial m}{\partial X} = d_1 X^{d_1 - 1} Y^{d_2} Z^{d_3}$$

which is either zero or of degree $d-1$. As f is homogeneous it is the sum of monomials of the same degree d hence $\frac{\partial m}{\partial X}$ is a sum of monomials of degree $d-1$. The same is of course true for the other partial derivatives.

Now

$$X \cdot \frac{\partial m}{\partial X} = X d_1 X^{d_1 - 1} Y^{d_2} Z^{d_3} = d_1 X^{d_1} Y^{d_2} Z^{d_3} = d_1 \cdot m\,.$$

Correspondingly,

$$Y \cdot \frac{\partial m}{\partial Y} = d_2 \cdot m, \qquad Z \cdot \frac{\partial m}{\partial Z} = d_3 \cdot m\,.$$

In particular

$$X \cdot \frac{\partial f}{\partial X} + Y \cdot \frac{\partial f}{\partial Y} + Z \cdot \frac{\partial f}{\partial Z} = (d_1 + d_2 + d_3) \cdot m = d \cdot m\,.$$

As f is the sum of monomials of the same degree and the differentiation is linear we obtain the result. □

With the information from this proposition we know that Equation (4.1) makes sense as the partial differentiated polynomials are homogeneous. Also, in case that $\operatorname{char} \mathbb{K} \nmid d$ (e.g. $\operatorname{char} \mathbb{K} = 0$) we do not need the condition that P lies on the curve ($P \in \mathfrak{V}(f)$, i.e. $f(P) = 0$) in the definition. as by Equation (4.2) from

$$\frac{\partial f}{\partial X}(P) = \frac{\partial f}{\partial Y}(P) = \frac{\partial f}{\partial Z}(P) = 0$$

it follows that $d \cdot f(P) = 0$. Hence, as $d \neq 0$ the relation $f(P) = 0$ is automatic.

Finally, we want to relate the singularities of an affine curve with the singularities of its projective completion. Let

$$i : \mathbb{A}^2(\mathbb{K}) \longrightarrow \mathbb{P}^2(\mathbb{K})$$

be the embedding of the affine plane into the projective plane given by

$$(\alpha, \beta) \longmapsto i(\alpha, \beta) := (\alpha : \beta : 1).$$

Let $Q = (\alpha, \beta)$ and $i(Q) = P = (\alpha : \beta : 1)$. Furthermore, let $f \in \mathbb{K}[X, Y, Z]$ be the homogenization of the polynomial $g \in \mathbb{K}[X, Y]$. In particular, $g(X, Y) = f(X, Y, 1)$.

Proposition 4.8. *The point Q is a singular point of $\mathfrak{V}(g) \iff i(Q)$ is a singular point for $\mathfrak{V}(f)$.*

Proof. We check the Euler Relation (4.2) for the point $i(Q) = (\alpha : \beta : 1)$

$$d \cdot f(\alpha, \beta, 1) = \alpha \cdot \frac{\partial f}{\partial X}(\alpha, \beta, 1) + \beta \cdot \frac{\partial f}{\partial Y}(\alpha, \beta, 1) + 1 \cdot \frac{\partial f}{\partial Z}(\alpha, \beta, 1)\,.$$

This can be rewritten as

$$d \cdot g(a, b) = \alpha \cdot \frac{\partial g}{\partial X}(\alpha, \beta) + \beta \cdot \frac{\partial g}{\partial Y}(\alpha, \beta) + 1 \cdot \frac{\partial f}{\partial Z}(i(Q))\,. \tag{4.3}$$

Let $Q = (\alpha, \beta)$ be a singular point of the affine curve $\mathfrak{V}(g)$. Then

$$\frac{\partial g}{\partial X}(Q) = \frac{\partial g}{\partial Y}(Q) = g(Q) = 0\,.$$

If we plug these into Equation (4.3) and use $\frac{\partial f}{\partial X}(\alpha, \beta, 1) = \frac{\partial g}{\partial X}(\alpha, \beta)$ and $\frac{\partial f}{\partial Y}(\alpha, \beta, 1) = \frac{\partial g}{\partial Y}(\alpha, \beta)$ we obtain

$$0 = 0 + 0 + \frac{\partial f}{\partial Z}(i(Q))\,.$$

This implies $\frac{\partial f}{\partial Z}(i(Q)) = 0$. The relations

$$\frac{\partial f}{\partial X}(i(Q)) = \frac{\partial g}{\partial X}(Q) = 0, \quad \frac{\partial f}{\partial Y}(i(Q)) = \frac{\partial g}{\partial Y}(Q) = 0\,,$$

and $f(i(Q)) = g(P) = 0$ are automatic. Hence $i(Q)$ is a singular point of the projective curve $\mathfrak{V}(f)$.

Vice versa, if $i(Q)$ is a singular point of $\mathfrak{V}(f)$ then all partial derivative are zero, and $g(Q) = f(i(Q)) = 0$. Hence Q is a singular point of the affine curve $\mathfrak{V}(g)$.

□

This has as a consequence that for a projective curve we can check whether its affine points are singular points also by checking the affine part of the curve. In addition we have to consider only the points of the curve which are intersection points with the line at infinity (i.e. have the third coordinate $\gamma = 0$) . This will be our strategy in the following.

Chapter 5

Projective lines and quadrics

Recall that $C = \mathfrak{V}(f)$ with $f \in \mathbb{K}[X, Y, Z]$, $\deg f > 1$ is a curve in the plane. We assume that f does not have multiple irreducible factors and defined $\deg C = \deg f$. The curve C is defined over the field $\mathbb{K}$. If $\mathbb{L} \geq \mathbb{K}$ is an field extension then $\mathfrak{V}(f)[\mathbb{L}]$ is called the set of $\mathbb{L}$-valued points of $\mathfrak{V}(f)$.

5.1 Lines

In this section we consider projective lines. This means $\deg C = \deg f = 1$. This implies that $C = \mathfrak{V}(f)$ with

$$f(X, Y, Z) = a \cdot X + b \cdot Y + c \cdot Z, \qquad a, b, c \in \mathbb{K} \tag{5.1}$$

and at least one of the a, b, c is $\neq 0$.

The curve C is given as

$$C = \{(\alpha, \beta, \gamma) \in \mathbb{P}^2(\mathbb{K}) \mid f(\alpha, \beta, \gamma) = 0\}.$$

For $\mathbb{C}_\infty$, the "line at ∞" with respect to the affine subset U_2, we have the defining polynomial $f_\infty(X, Y, Z) = Z$ and

$$C_\infty = \{(\alpha, \beta, 0) \in \mathbb{P}^2(\mathbb{K}) \mid \alpha, \beta \in \mathbb{K}, \alpha \neq 0 \text{ or } \beta \neq 0\}.$$

Proposition 5.1. *1. A line is always nonsingular.*

2. Let L_1,L_2 be two lines, $L_1 \neq L_2$ then $\exists$ a unique point P with $\{P\} = L_1 \cap L_2$. This says that two lines always have a unique intersection.

3. Given two points P_1,P_2 with $P_1 \neq P_2$ then there exists a unique line L which passes through P_1 and P_2, i.e. P_1,$P_2 \in L$.

Proof. **1.** Let $L = \mathfrak{V}(f)$, with f given by Equation (5.1) . We calculate the derivatives

$$\frac{\partial f}{\partial X} = a\,, \quad \frac{\partial f}{\partial Y} = b\,, \quad \frac{\partial f}{\partial X} = c\,.$$

Hence at least one of them is different from zero. This implies that there does not exist singular points.

2. Given the two lines as

$$L_1 : \mathfrak{V}(a_1 \cdot X + b_1 \cdot Y + c_1 \cdot Z)\,, \qquad L_2 : \mathfrak{V}(a_2 \cdot X + b_2 \cdot Y + c_2 \cdot Z)\,.$$

The condition $L_1 \neq L_2$ implies that

$$\forall \lambda \in \mathbb{K} : (a_1, b_1, c_1) \neq \lambda \cdot (a_2, b_2, c_2)\,. \tag{5.2}$$

A point $P = (\alpha : \beta : \gamma) \in L_1 \cap L_2$ if and only if

$$a_1 \cdot \alpha + b_1 \cdot \beta + c_1 \cdot \gamma = 0$$

$$a_2 \cdot \alpha + b_2 \cdot \beta + c_2 \cdot \gamma = 0\,.$$

This is a homogeneous system of linear equations. Equation (5.2) implies that the rank of the system equals 2. Hence the space of solutions (in $\mathbb{K}^3$) is one-dimensional, i.e.

$$\{\lambda \cdot \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}\}\,,$$

with a vector $(\alpha, \beta, \gamma) \neq \emptyset$. Interpreted in the projective plane this corresponds to just a unique point $P = (\alpha : \beta : \gamma)$. This was the claim.

3. This is the dual problem. We have two points $P_1 = (\alpha_1 : \beta_1 : \gamma_1)$, $P_2 = (\alpha_2 : \beta_2 : \gamma_2)$. As $P_1 \neq P_2$ the set of homogeneous coordinates are not multiples of each others. We search a polynomial

$$f(X, Y, Z) = a \cdot X + b \cdot Y + c \cdot Z$$

such that

$$f(\alpha_1, \beta_1, \gamma_1) = 0, \qquad f(\alpha_2, \beta_2, \gamma_2) = 0\ .$$

This is a system of 2 linear equations of rank 2. Hence again the space of solution is one-dimensional and the function f is fixed up to multiplication with a non-zero constant. This says $C = \mathfrak{V}(f)$ is uniquely given. □

5.2 Quadrics in $\mathbb{P}^2(\mathbb{K})$

Quadrics are curves of degree two. Hence the degree of the defining polynomial f is two. In its most general form the polynomial can be written as

$$f(X,Y,Z) = a \cdot X^2 + b \cdot Y^2 + c \cdot Z^2 + d \cdot XY + e \cdot XZ + h \cdot YZ, \tag{5.3}$$

with $a, b, c, d, e, h \in \mathbb{K}$.

We now make a change of projective coordinates to simplify this expression (For more details about change of coordinates see the next chapter). It will turn out that $\operatorname{char}\mathbb{K} = 2$ will need a special treatment which we are not intend to do here. Hence we assume $\operatorname{char}\mathbb{K} \neq 2$ in the following. Our goal is to remove the mixed terms in (5.3) .

Let us first assume $a \neq 0$. We express

$$a(X + \frac{d}{2a}Y + \frac{e}{2a}Z)^2 = a \cdot X^2 + d \cdot XY + e \cdot XZ + \frac{de}{4a}\, YZ + \frac{d^2}{4a}\, Y^2 + \frac{e^2}{4a}\, Z^2 .$$

(*Question:Do you see why we needed to exclude* $\operatorname{char}\mathbb{K} = 2$ *?*)

We plug this expression into the expression (5.3) for f and introduce new variables

$$X' = (X + \frac{d}{2a}Y + \frac{e}{2a}Z)$$
$$Y' = Y$$
$$Z' = Z .$$

Our polynomial f can be written in the new coordinates as

$$a \cdot (X')^2 + \tilde{b} \cdot (Y')^2 + \tilde{c} \cdot (Z')^2 + \tilde{h} \cdot Y'Z' ,$$

with suitable (calculable) $\tilde{b}$, $\tilde{c}$, and $\tilde{d} \in \mathbb{K}$.

We obtain less terms. In particular the X' appears purely quadratic. Assuming that $\tilde{b}$ (or $\tilde{c}$)$\neq 0$ we repeat the step above accordingly to remove the mixed term $Y'Z'$.

If $a = 0$ we check whether b or c are different from zero and start with the corresponding variable. But we are stuck if we only have mixed terms. For example XY . In this case we rewrite

$$X \cdot Y = \frac{1}{4}((X+Y)^2 - (X-Y)^2)$$

and set

$$X' = X + Y \qquad Y' = X - Y \qquad Z' = Z$$

(Note that $X = \frac{1}{2}(X' + Y')$).

In any case we will obtain an expression of the type

$$a \cdot (X')^2 + \text{rest}$$

and can continue as above. Finally we obtain

Proposition 5.2. *Over a field* $\mathbb{K}$ *of* $\operatorname{char}\mathbb{K} \neq 2$ *each quadric can by given with respect to suitable coordinates as* $C = \mathfrak{V}(f)$ *with the polynomial*

$$f(X, Y, Z) = a \cdot X^2 + b \cdot Y^2 + c \cdot Z^2\,, \qquad (a, b, c) \neq (0, 0, 0)\,.$$

Attention: Some of the a,b,c can be zero but not all of them together!

Now assume that $\mathbb{K}$ is an algebraically closed field. In this case we go one step further. Assume e.g. $a \neq 0$ then we take as new coordinate $X' = (\sqrt{a}X)$ and get rid of the coefficient a. Hence we obtain the following 3 cases (up to change of coordinates)

(1) $X^2 + Y^2 + Z^2$

(2) $X^2 + Y^2$

(3) X^2

Case (1) : We look for singularities

$$\frac{\partial f}{\partial X} = 2X, \quad \frac{\partial f}{\partial Y} = 2Y, \quad \frac{\partial f}{\partial Z} = 2Z\,.$$

The point $(\alpha : \beta : \gamma)$ is a simultaneous zero if and only if $\alpha = \beta = \gamma = 0$; But this is not a point in the projective plane. Hence

Proposition 5.3. $C = \mathfrak{V}(X^2 + Y^2 + Z^2)$ *is a nonsingular quadric.*

Case (2) : As $\mathbb{K}$ is algebraically closed, there is a root $\sqrt{-1}$ of the polynomial $X^2 + 1$. In particular $\sqrt{-1} \cdot \sqrt{-1} = -1$. Now we calculate

$$X^2 + Y^2 = (X + \sqrt{-1} \cdot Y) \cdot (X - \sqrt{-1} \cdot Y)\,.$$

Hence, our f decompose into two linear polynomials. Consequently, the quadric C is the union of two different lines. The singular points are the simultaneous zeros of

$$\frac{\partial f}{\partial X} = 2X, \quad \frac{\partial f}{\partial Y} = 2Y, \quad \frac{\partial f}{\partial Z} = 0\,.$$

Hence $\alpha = \beta = 0$ and γ arbitrary. We get one singular point $(0:0:1)$. This is an affine point with respect to U_2. Its affine coordinate is $(0,0)$.

Case (3) : In this case

$$C = \mathfrak{V}(X^2) = \mathfrak{V}(X) = \{(0:\beta:\gamma) \mid \beta \neq 0 \quad \text{or} \quad \gamma \neq 0\}$$

This is just one straight line given by the Y-axis. Strictly speaking, from a higher point of view, we should consider C as $\mathfrak{V}(X^2)$ as a line with multiplicity 2, i.e. a double line. From this description we would obtain also that all points are singular points.

In this way we obtained a complete classification of quadrics over algebraically closed fields.

If $\mathbb{K}$ is not algebraically closed the situation is much more complicated. As an example we give a few remarks for $\mathbb{K} = \mathbb{R}$. If $a > 0$ we can argue as above. If $a < 0$ then we compute it as $a = -(-a)$ with $-a > 0$. Consequently we obtain the following different cases:

(1a) $X^2 + Y^2 + Z^2$

(1b) $X^2 + Y^2 - Z^2$ (which is the same as $X^2 - Y^2 - Z^2$)

(2a) $X^2 + Y^2$

(2b) $X^2 - Y^2$

(3) X^2

We will only have a look on (1a) and (1b). Clearly these curves will be nonsingular, as they do not have singular points over the algebraic closure. Note that by the passage to $\overline{\mathbb{R}} = \mathbb{C}$ both cases (1a) and (1b) will yield the case (1) above.

In case (1a) : We get $\mathfrak{V}(f) = \emptyset$.

In case (1b) : We have solutions

$$C = \{(\alpha:\beta:\gamma) \mid \alpha^2 + \beta^2 - \gamma^2 = 0\}\,.$$

What does the set of solutions look like? For this we want to study them on affine parts.

1. First we consider the solution set C on the affine part given by $\gamma = 1$ (this means on U_2) then
$$C^{(1)}_{\text{aff}} := \{(\alpha, \beta) \mid \alpha^2 + \beta^2 = 1\},$$
hence it is a circle. This does not have a "point at ∞" as $\gamma = 0$ and $\alpha^2+\beta^2 = 0$ implies $\alpha = \beta = \gamma = 0$.

2. Next we consider points given by the affine part $\beta = 1$ (i.e. U_1). Then
$$C^{(2)}_{\text{aff}} := \{(\alpha, \gamma) \mid \alpha^2 + 1 - \gamma^2 = 0\} = \{(\alpha, \gamma) \mid \gamma^2 = \alpha^2 + 1\}.$$
This is a hyperbola. The points at ∞ are given by $\beta = 0$. This implies $\alpha^2 - \gamma^2 = 0 = (\alpha + \gamma)(\alpha - \gamma))$. Hence there are two points $(1 : 0 : 1)$ and $(1 : 0 : -1)$.

3. The affine part with respect to U_0, corresponds to the previous situation. But we could also take other affine parts. Recall that $X^2 + Y^2 - Z^2$ corresponds to $X^2 + YZ$ after a change of coordinates.

   Now we take $\gamma = 1$ in the new coordinates and get
$$C^{(3)}_{\text{aff}} := \{(\alpha, \beta) \mid \beta = -\alpha^2\},$$
which is a parabola. There is one point at ∞. It corresponds to $\alpha = 0$ and yields the point $(0 : 1 : 0)$.

With the above we saw that from the real projective point of view all the objects: circles, hyperbolas, and parabolas are different affine parts of one projective object, the real projective quadric $\mathfrak{V}(X^2 + Y^2 - Z^2)$.

Chapter 6

Transformation of variables

We already dealt with the transformation of variables in the previous chapter. Now we want to have a closer look on them. After we fix an origin O and a basis in $\mathbb{K}^n$, we have the identification $\mathbb{A}^n(\mathbb{K}) \cong \mathbb{K}^n$. Also

$$\mathbb{P}^n(\mathbb{K}) \cong (\mathbb{K}^{n+1} \backslash \{0\}) / \sim \quad ,$$

where the identification is given after choosing a basis in $\mathbb{K}^{n+1}$.

6.1 Affine transformations

The following is just linear algebra. Let V be a vector space of $\dim V = n$, and B and B' two sets of basis elements of V. We denote $B = (e_1, ..., e_n)$ and $B' = (e'_1, ..., e'_n)$.

The base transformation matrix $M := M_{B \to B'}$ from the old basis B to the new basis B' is the invertible $n \times n$ matrix which has in the j^{th} column (M_{*j}) the coefficient of the new basis vector e'_j in B' with respect to the old basis vectors in B.

Let P be a vector (respectively point) and $\alpha \in \mathbb{K}^n$ its coordinate vector with respect to the old basis B and α' with respect to the new basis. Then we have the relations:

$$\alpha = M \cdot \alpha' \,, \quad \alpha' = (M)^{-1} \cdot \alpha \,.$$

We set $N = M^{-1}$.

This is the situation for a vector space. For the affine space we have also to take a translation of the origin O into account.

Hence: an affine transformation ϕ is given by

$$\alpha' = N \cdot \alpha + t\,, \quad t \in \mathbb{K}^n\,.$$

The map ϕ is uniquely fixed by the pair $(N,t)\,, N \in GL(n,\mathbb{K})\,, t \in \mathbb{K}^n\,.$

Exercise: Write the formula for the composition of affine transformations. Give the expression for the inverse transformation.

6.2 Projective transformations

Here the origin is always fixed by $0 \in \mathbb{K}^{n+1}\,.$ Hence with $\alpha \in \mathbb{K}^{n+1}$

$$\alpha' = N \cdot \alpha\,, \quad N \in GL(n+1,\mathbb{K})\,.$$

Recall that for the points $[\alpha] = [\beta]$ in projective space we have

$$\alpha \sim \beta \leftrightarrow \beta = \lambda \cdot \alpha\,, \lambda \in \mathbb{K}^*\,.$$

Hence if $N = \lambda \cdot I_{n+1}\,, \lambda \neq 0$ the corresponding projective transformation ϕ is the identity :

$$N' = \lambda \cdot N \quad \leftrightarrow \quad \phi_N = \phi'_N\,.$$

The set $\{\lambda \cdot I_{n+1} \mid \lambda \in \mathbb{K}^*\}$ with I_{n+1} the $(n+1)\times(n+1)$ identity matrix, constitutes a normal subgroup of $GL(n+1,\mathbb{K})\,,$ and the projective transformations are given by the elements of the quotient group

$$PGL(n+1,\mathbb{K}) := GL(n+1,\mathbb{K})/\{\lambda \cdot I_{n+1} \mid \lambda \in \mathbb{K}^*\}\,.$$

This group is called projective linear group.

6.3 Transformation of affine and projective varieties

One has two different ways of interpretations of the above transformations:

- change of coordinates

- action on the ambient space $\mathbb{K}^n\,,$ respectively on $\mathbb{K}^{n+1}\,.$

Take $\underline{X} = (X_1, X_2, ..., X_n)$ and $f(\underline{X}) \in \mathbb{K}[X]$ a polynomial. Given an affine transformation (N, t) we set

$$f_{N,t}(\underline{X}) := f(N \cdot \underline{X} + t)\,.$$

Note that in general $f_{N,t} \neq f$. Given a variety $V = \mathfrak{V}(f_1, ..., f_k)$, we set $V_{N,t} = \mathfrak{V}(f_{1,N,t}, ..., f_{k,N,t})$. In general both varieties will be different, but the points in the two sets will be in $1:1$ correspondence under an affine transformation.

Definition 6.1. Let V and V' be two affine varieties. V is called *affine equivalent* to V' if and only if there exists an $N \in GL(n, \mathbb{K})$ and $t \in \mathbb{K}^n$ such that $V' = V_{N,t}$.

Obviously, this is an equivalence relation. Geometrically V' is obtained from V by an affine map

$$\mathbb{A}^n(\mathbb{K}) \longrightarrow \mathbb{A}^n(\mathbb{K})\,.$$

What has been done for affine variety makes sense also for projective varieties.

Definition 6.2. Let V and V' be two projective varieties, then V and V' are called projectively equivalent if and only if $\exists N \in GL(n+1, \mathbb{K})$ such that $V' = V_N$.

Exercise : Let V and V' be two affine plane curves which are affine equivalent then their projective completions are projectively equivalent (Attention: the opposite direction is not true).

We give the relation between the affine and the projective transformations. Recall that the affine transformation is given by

$$\begin{pmatrix} X' \\ Y' \end{pmatrix} = A \cdot \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} t_X \\ t_Y \end{pmatrix}\,, \quad A = (a_{ij})$$

respectively

$$X' = a_{11}X + a_{12}Y + t_X$$

$$Y' = a_{21}X + a_{22}Y + t_Y\,.$$

We make this homogeneous via

$$t_X \longmapsto t_X Z \qquad t_Y \longmapsto t_Y Z\,,$$

and the transformation reads as

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \left(\begin{array}{ccc|c} & A & & t_X \\ & & & t_Y \\ \hline 0 & & 0 & 1 \end{array}\right) \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} . \tag{6.1}$$

In this sense for the group of affine respectively projective transformations we have

$$G(\mathbb{A}^n(\mathbb{K})) \quad \lneqq \quad G(\mathbb{P}^n(\mathbb{K}))$$

as the last line of the projectivized matrix is always $(0, 0, 1)$. In particular, there are more projective maps than affine maps.

Example. Let $\mathbb{K} = \mathbb{R}$. We discussed already earlier the 2 curves,

$$V_1 : X^2 + Y^2 - 1 \qquad V_2 : X^2 - Y^2 - 1 .$$

V_1 and V_2 are not equivalent as affine curves. Recall that V_1 is a circle and V_2 is a hyperbola.

If we homogenize the defining equations we obtain

$$\tilde{V_1} : X^2 + Y^2 - Z^2 \qquad \tilde{V_2} : X^2 - Y^2 - Z^2 \quad (= -X^2 + Y^2 + Z^2) .$$

The projective transformation relating both $\tilde{V_1}$ and $\tilde{V_2}$ is given as interchanging X and Z

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} .$$

Hence $\tilde{V_1} \sim \tilde{V_2}$. Note that this matrix is not of the type (6.1). Hence, the projective transformation does not come from an affine transformation.

6.4 Singularities and intersections

Proposition 6.3. *1. Let $\Phi : \mathbb{P}^2 \longrightarrow \mathbb{P}^2$ be a transformation. Then $P \in \mathfrak{V}(f)$ is a singular point of C if and only if $\Phi(P)$ is a singular point of $C_\Phi = \mathfrak{V}(f_\Phi)$.*

2. The definition of a singular point does not depend on the basis chosen.

Proof. **1.** Let $\underline{X'} = \Phi(\underline{X}) = A \cdot \underline{X}$ be the projective transformation associated with the invertible matrix A. If P is a point on C with homogeneous coordinates α then $\Phi(P)$ is a point on C_Φ with homogeneous coordinates $\alpha' = A^{-1}\alpha$. Furthermore,

$$f_\Phi(\underline{X}) = f'(A \cdot \underline{X}) = f'(\underline{X'})\,.$$

Hence by the chain rule

$$\frac{\partial f'}{\partial \underline{X'}} = \frac{\partial f_\Phi}{\partial \underline{X'}} = A^{-1} \cdot \frac{\partial f}{\partial \underline{X}} \quad , \quad A = \frac{\partial \Phi}{\partial \underline{X}}\,.$$

A and A^{-1} have full rank 3, hence

$$\frac{\partial f}{\partial \underline{X}}(P) = 0 \quad \Longleftrightarrow \quad \frac{\partial f'}{\partial \underline{X'}}(\Phi(P)) = 0\,.$$

2. follows directly from (a), as a change of basis corresponds to such a Φ. □

Important consequence: We can calculate the singular points in suitable coordinates.

Proposition 6.4. *Let $L = \mathfrak{V}(aX + bY + cZ)$ be a projective line then there exists a projective transformation Φ such that $L' = \mathfrak{V}(Z') = L_\Phi \cong L$.*

Geometrically this means that in the projective plane we can move a given projective line "to infinity".

Proof. Case 1: If $c \neq 0$ we take $Z' = aX + bY + cZ\,,\ X' = X\,,\ Y' = Y$. The matrix for the transformation reads as

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & b & c \end{pmatrix}\,.$$

Case 2,: If $c = 0$ then at least one of a or b are $\neq 0$. For example if $a \neq 0$ we set $Z' = aX + bY\,,\ X' = Z\,,\ Y' = Y$ and obtain

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ a & b & 0 \end{pmatrix}$$

and if $b \neq 0$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ a & b & 0 \end{pmatrix} .$$

□

The following theorem will be of geometric importance later on.

Theorem 6.5. *Let C be a projective curve of degree n and L be a projective line. Assume L is not a component of C then $C \cap L$ is a set of finitely many points and*

$$\#(C \cap L) \leq n \, .$$

Proof. Let L be the projective line. Following Proposition 6.4 we can assume after a change of projective coordinates that $L = \mathfrak{V}(Z)$. Let $C = \mathfrak{V}(f)$. Note that

$$(\alpha : \beta : \gamma) \in L \quad \leftrightarrow \quad \gamma = 0 \, .$$

Hence $(\alpha : \beta : \gamma) \in C \cap L$ if and only if $f(\alpha : \beta : 0) = 0$. We set $g(X, Y) = f(X, Y, 0)$, and f is homogeneous of degree n, i.e. it is a sum of monomials $X^{d_1} Y^{d_2} Z^{d_3}$ with $d_1 + d_2 + d_3 = n$. Only those monomials will survive for g for which $d_3 = 0$. But this means that g will be either identically zero or it will be homogeneous of degree n (now of two variables).

If $g \equiv 0$ then $L \subset \mathfrak{V}(f)$. This says L is a component of C. But this was excluded. Hence, our g can be written as

$$g(X, Y) = a_0 Y^n + a_1 X Y^{n-1} + ... + a_{n-1} X^{n-1} Y + a_n X^n \, ,$$

where not all $a_k = 0$. If $\alpha = (1 : 0 : 0) \in L \cap C$ then $g(1, 0) = 0 = a_n \cdot 1$. This implies $a_n = 0$. If $\alpha \in L \cap C$ but $\alpha \neq (1 : 0 : 0)$ then we can write $\alpha = (x : 1 : 0)$,

$$g(\alpha) = \hat{g}(x) = a_0 + a_1 x + ... + a_{n-1} x^{n-1} + a_n x^n \, .$$

Hence we have maximally n zeros of this type. If $(1 : 0 : 0) \in L \cap C$ then as $a_n = 0$ we will have maximally $(n - 1)$ roots of this type. Always the number of intersection points will be bounded by n. □

Additional comments:

1. We can add intersection multiplicities for the bound (see later).
2. For $\mathbb{K} = \bar{\mathbb{K}}$ and counted with multiplicities the above proof shows

$$\#C \cap L = n\,.$$

In fact the above theorem has an extension.

Theorem 6.6. *(Theorem of Bezout). Let C_1 and C_2 be two projective (planar) curves defined over $\mathbb{K}$ without common components in the algebraic closure $\overline{\mathbb{K}}$ then*

$$\sum_{P \in C_1 \cap C_2} m(P, C_1, C_2) \leq (\deg C_1) \cdot (\deg C_2). \tag{6.2}$$

If the field $\mathbb{K}$ is already algebraically closed, then in (6.2) we have equality.

We will not prove it here. Above we proved it if one of the curves is a line. For a correct interpretation in general we would first have to extend the definition of the multiplicity to $m(P, C_1, C_2)$, which is not so simple. In case that the intersection point P is a non-singular point on both curves and that the tangent lines of both curves at this point P are distinct the multiplicity will be one (in complete accordance with the case $m(P, C, L)$ as in this case L is its own tangent and different tangents means that L is not a tangent at the point P to the curve C).

Chapter 7

Elliptic curves

7.1 Basic definitions

Next we consider projective curves C of degree 3, this says

$$C = \mathfrak{V}(f)[\mathbb{K}], \qquad f(X,Y,Z) \in \mathbb{K}[X,Y,Z], \quad \text{with} \quad \deg f = 3\,.$$

Definition 7.1. A curve $C \subseteq \mathbb{P}^2(\mathbb{K})$ is called an elliptic curve if and only if C is a non-singular projective curve of $\deg C = 3$.

In case that $\deg C = 3$ (without requiring that C is nonsingular) we call C a cubic curve.

Example. Let $f = l_1 \cdot l_2 \cdot l_3$ with l_i linear polynomials. Then $C = \mathfrak{V}(f)$ is a cubic curve. It is the union of 3 lines where the intersection points are singular points of the cubic curve. Hence it is not an elliptic curve (see Figure 7.1)

In fact our elliptic curves will always be indecomposable. This follows from the fact, that if C has e.g. two components over the algebraic closure then they will have a point of intersection (over the algebraic closure) which will be a singular point. Recall that "nonsingular curve" means "no singular points over the algebraic closure $\overline{\mathbb{K}}$".

A general cubic homogeneous polynomial of degree 3 in 3 variables is a linear combination of 10 different monomials, i.e. it has 10 parameters. But by a change of variables we obtain

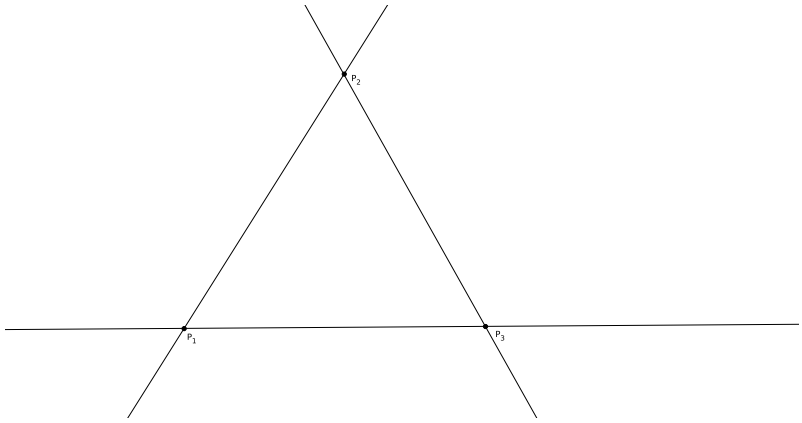

Figure 7.1: Three lines

Proposition 7.2. *Let E be an elliptic (projective) curve which has (at least) one $\mathbb{K}$-valued point then after a change of variables the curve can be given as $E = \mathfrak{V}(f)$ with*

$$f(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - (X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3)\,, \quad (7.1)$$

with $a_1, \ldots, a_6 \in \mathbb{K}$.

Proposition 7.3. (a) *If $\mathrm{char}(\mathbb{K}) \neq 2$ then in (7.1) $a_1 = a_3 = 0$ can be obtained by a further change of variables.*

(b) *If $\mathrm{char}(\mathbb{K}) \neq 2, 3$ then in (7.1) even $a_1 = a_3 = a_2 = 0$ can be obtained. In particular (7.1) transforms to*

$$Y^2Z - (X^3 + a_4XZ^2 + a_6Z^3)\,. \quad (7.2)$$

The expression (7.2) is called *Weierstraß normal form*. We will not carry through the proofs, you might consult the book [Wer, p.24-27].

Attention: We do not claim that a curve C given in the form of (7.2) or (7.1) will be automatically nonsingular (see below).

Points at "∞": (with respect to standard affine coordinates $(\alpha : \beta : 1)$).

For this we have to put $Z = 0$ (i.e. $\gamma = 0$) in our defining equation (7.1). It remains the equation $-X^3$, hence $\alpha = 0$ and β arbitrary. This says there is a unique

point $Q = (0 : 1 : 0)$ lying on the elliptic curve and on the line at ∞, $L = \mathfrak{V}(Z)$. This is the point

$$O := L \cap C = \{(0 : 1 : 0)\}\,.$$

We will discuss later that it is an intersection point of multiplicity 3.

Singular points: The cases $\operatorname{char}\mathbb{K} = 2$ or 3 would need a special treatment case by case which we will not do here.[1] In the other cases we can start from the Weierstraß normal form

$$f(X, Y, Z) = -Y^2Z + (X^3 + a_4XZ^2 + a_6Z^3)\,.$$

Singular points will be the simultaneous zeros of the partial derivatives of f (as the characteristic of $\mathbb{K}$ does not divide 3)

$$\begin{aligned}
\frac{\partial f}{\partial X}(X, Y, Z) &= 3X^2 + a_4Z^2\,,\\
\frac{\partial f}{\partial Y}(X, Y, Z) &= -2YZ\,,\\
\frac{\partial f}{\partial Z}(X, Y, Z) &= -Y^2 + 2a_4XZ + 3a_6Z^2\,.
\end{aligned}$$

First we check the point $(0 : 1 : 0)$ at ∞, whether it is a singular point or not. But $\frac{\partial f}{\partial Z}(0, 1, 0) = -1^2 \neq 0$, hence it is not a singular point.

Now we consider the affine part of E. For Z we plug in 1 and obtain from above for the point with affine coordinates (x, y) directly $y = 0$ (2nd equation). It remains to check the other two equations. Let us assume that we have a singular point $(x, y) = (x, 0)$. Then

$$3x^2 + a_4 = 0 \to 3x^2 = -a_4\,,$$

$$2a_4x + 3a_6 = 0 \to 2a_4x = -3a_6\,.$$

Now we multiply the first equation by $4a_4^2$, take the square of the 2nd equation, multiply it with 3 (note that $\operatorname{char}\mathbb{K} = 2, 3$ was excluded) and we obtain

$$12a_4^2x^2 = -4a_4^3\,,$$

$$12a_4^2x^2 = 27a_6^2\,.$$

Hence under the assumption that there is a singular point we obtain,

$$\Delta = (4a_4^3 + 27a_6) = 0\,. \tag{7.3}$$

[1] Of course, the reader is invited to do it himself/herself.

The expression Δ is called discriminant. Up to now we know that if $\Delta \neq 0$ there will be no singular points, hence $E = \mathfrak{V}(f)$ will be an elliptic curve. For the opposite (meaning if there is a singular point then $\Delta = 0$), it is enough to consider the affine part, as $(0:1:0)$ is not a singular point, hence the points fulfilling

$$y^2 = g(x)\,. \quad \text{with} \quad g(X) = X^3 + a_4X + a_6 \in \mathbb{K}[X]. \tag{7.4}$$

We pass over to the algebraic closure $\overline{\mathbb{K}}$, over $\mathbb{K}$ and decompose the degree 3 polynomial g into 3 linear polynomials as factors (we write it as an equation, with $e_1, e_2, e_3 \in \overline{\mathbb{K}}$)

$$\begin{aligned} g(X) &= (X-e_1)(X-e_2)(X-e_3) \\ &= X^3 - (e_1+e_2+e_3)X^2 + (e_1\cdot e_2 + e_1\cdot e_2 + e_2\cdot e_3)X - e_1\cdot e_2\cdot e_3\,. \end{aligned}$$

If we compare the coefficients from (7.4) with the coefficients above we obtain for the discriminant (7.3)

$$\Delta = A\cdot(e_1-e_2)^2\cdot(e_1-e_3)^2\cdot(e_2-e_3)^2\,.$$

Here A is a constant $\neq 0$. Hence Δ is the square of the differences of the zeros of the polynomial g. (This is the usual algebraic discriminant of polynomials).

As a consequence we obtain that $\Delta = 0$ implies that at least two of the e_i have to coincide. We now show that then there is a singular point (x, y).

If two are the same then, e.g. $e_1 = e_2$ then

$$f(X,Y) = -Y^2 + (X-e_1)^2(X-e_3)\,. \tag{7.5}$$

First of course $\frac{\partial f}{\partial Y} = 0$ implies as above $y = 0$. Then

$$\begin{aligned} \frac{\partial f}{\partial X}(x,0) = \frac{\partial g}{\partial X}(x) &= 2\cdot(x-e_1)(x-e_3) + (x-e_1)^2 \\ &= (x-e_1)(3x-2e_3-e_1)\,. \end{aligned}$$

Two candidates for singular points exists. First we consider $x = e_1$. In this case $(e_1, 0)$ lies on the curve E. Hence it is a singular point. The other point with $(x = \frac{1}{3}(2e_3 + e_1), 0)$ does not lie on the curve (see (7.5)).

Hence there is a unique singular point $(e_1, 0)$, respectively $(e_1 : 0 : 1)$. The corresponding cubic curve is called nodal cubic, see Figure 3.2.

If all three e_i are the same then

$$\frac{\partial f}{\partial X}(x,0) = \frac{\partial g}{\partial X}(x) = 3\cdot(x-e_1)^2\,.$$

Hence $(e_1, 0)$ is a singular point as obviously it lies on the curve. As $e_1+e_2+e_3 = 0$ this implies $e_1 = e_2 = e_3 = 0$. In this case we obtain the cuspidal cubic, see Figure 3.1 .

We have shown

Theorem 7.4. (a) *Let* $\operatorname{char}\mathbb{K} \neq 2, 3$ *and denote by* $E = E(a_4, a_6)$ *the curve given by the polynomial*

$$Y^2Z - (X^3 + a_2XZ^2 + a_4Z^3)\,,$$

then E *is nonsingular (i.e. elliptic) if and only if*

$$\Delta = (4a_4^3 + 27a_6^2) \neq 0\,.$$

(b) *Let* $\overline{\mathbb{K}}$ *be the algebraic closure of* $\mathbb{K}$ *then the curve* E *can be given as*

$$Y^2Z - (X - e_1Z)(X - e_2Z)(X - e_3Z)\,, \quad e_i \in \overline{\mathbb{K}}\,.$$

The curve is nonsingular if and only if the e_1, e_2, e_3 *are pairwise distinct.*

Remark 7.5. (About the condition of existence of a $\mathbb{K}$-valued point in Proposition 7.2.) If we have the curve given in the form of (7.4) then the point at ∞, given by $(0 : 1 : 0)$, is always a $\mathbb{K}$-valued point. The a priori existence of a $\mathbb{K}$-valued point is needed before we can transform the defining polynomial into the normal form. If we work over an algebraically closed field we have always solutions as shown earlier, see Proposition 3.7. Hence, it is only a condition over algebraically non-closed fields.

7.2 Group structure of an elliptic curve

In the following let $E = \mathfrak{V}(f)$ be an arbitrary elliptic curve (i.e. a projective cubic curve which is nonsingular) defined over an arbitrary field $\mathbb{K}$. We assume that E has a $\mathbb{K}$-valued point. We will construct the structure of an abelian group for $\mathfrak{V}(f)[\mathbb{K}]$ or more general for $\mathfrak{V}(f)[\mathbb{L}]\,, \mathbb{L} \geq \mathbb{K}$, in a purely geometric manner.

To achieve this we will have to cut our elliptic curve with lines. For this to work correctly we have to study intersection multiplicities first.

Definition 7.6. Let $C = \mathfrak{V}(f)$ be a projective curve and $P \in C$ a nonsingular (i.e. a regular) point $P = (\alpha : \beta : \gamma) = \underline{\alpha}$. Then the line $L_p = \mathfrak{V}(l_p)$ defined via

$$l_p(X, Y, Z) = \frac{\partial f}{\partial X}(\underline{\alpha}) \cdot X + \frac{\partial f}{\partial Y}(\underline{\alpha}) \cdot Y + \frac{\partial f}{\partial Z}(\underline{\alpha}) \cdot Z\,, \tag{7.6}$$

is called *tangent line* of the curve C at the point $P \in C$.

Remark 7.7. 1. Tangent lines exists at regular points as in this case the equation (7.6) is not identically zero.

2. The point P lies itself on the tangent line at P, i.e. $P \in L_p$. Via the Euler relation

$$l_p(\underline{\alpha}) = \frac{\partial f}{\partial X}(\underline{\alpha}) \cdot \alpha + \frac{\partial f}{\partial Y}(\underline{\alpha}) \cdot \beta + \frac{\partial f}{\partial Z}(\underline{\alpha}) \cdot \gamma = d \cdot f(\underline{\alpha}) \,,$$

as $P \in C$ i.e. $f(\underline{\alpha}) = 0$, hence it follows that $l_p(\underline{\alpha}) = 0$, which says $P \in L_p$. In particular, $L_p \cap C \neq \emptyset$.

Next we consider arbitrary lines L and C a projective curve which is irreducible of degree $d > 1$. We already showed in Theorem 6.5 that $\#C \cap L \leq d$. Let $C = \mathfrak{V}(g)$ and $P \in L$ a fixed point $P = (\alpha : \beta : \gamma)$. Let $P' = (\alpha' : \beta' : \gamma') \neq P$ be another point on the line. We set

$$\psi(t) := g(\alpha + t\alpha', \beta + t\beta', \gamma + t\gamma') = g(P + tP') \,, \quad t \in \mathbb{K} \,. \tag{7.7}$$

The function $\psi(t)$ is a polynomial in t of degree $\leqq d$.

Definition 7.8. The order of the zero at $t = 0$ of the polynomial ψ is called *multiplicity of the point P in the intersection $L \cap C$*, in symbols $m = m(P, L, C)$.

1. Note that

$$\psi(t) = a_n t^n + a_{n+1} t^{n+1} + ... + a_d t^d = \sum_{k=n}^{d} a_k t^k \,,$$

with $a_n \neq 0$. If we take instead of P' another point P'' on the line then the coefficients may change but not the fact that n is the smallest order. Also the multiplicity will always be smaller than d.

2. If we put $t = 0$ in (7.7) then $\psi(0) = g(\alpha, \beta, \gamma)$ hence

$$\psi(0) = 0 \quad \leftrightarrow \quad P \in L \cap C \quad \leftrightarrow \quad m(P, L, C) > 0 \,.$$

3. We can also determine multiplicities of zeros of functions by considering derivatives. If we differentiate (7.7) with respect to the variable t we get

$$\psi'(t) = \frac{\partial g}{\partial X}(\underline{\alpha} + t\underline{\alpha}') \cdot \alpha' + \frac{\partial g}{\partial Y}(\underline{\alpha} + t\underline{\alpha}') \cdot \beta' + \frac{\partial g}{\partial Z}(\underline{\alpha} + t\underline{\alpha}') \cdot \gamma' \,.$$

Now $m \geq 2$ if and only if $\psi'(0) = 0$. This says

$$0 = \frac{\partial g}{\partial X}(\underline{\alpha}) \cdot \alpha' + \frac{\partial g}{\partial Y}(\underline{\alpha}) \cdot \beta' + \frac{\partial g}{\partial Z}(\underline{\alpha}) \cdot \gamma' .$$

This yields that the points $P' = (\alpha' : \beta' : \gamma')$ are lying on the tangent line of the curve C at the point P.

Proposition 7.9. *The multiplicity $m(P, L, C) \geq 2 \Leftrightarrow L$ is a tangent line along the curve at the point P.*

Now we return to the case that our curve is an elliptic curve E.

Theorem 7.10. *Let L be a projective line and E an elliptic curve then*

$$\sum_{P \in \mathbb{P}^2(\mathbb{K})} m(P, L, E) = \begin{cases} 0 \\ 1 \\ 3 \end{cases} .$$

First remark that $\sum_{P \in \mathbb{P}^2(\mathbb{K})}$ can be replaced by $\sum_{P \in L \cap E}$ as far as for the other points P we have $m(P, L, E) = 0$. Also note that $m(P, L, E) \leqq \deg E = 3$, by Theorem 6.5. In particular, the statement of the theorem is that a line will meet an elliptic curve counted with multiplicities either not at all, or only once or exactly 3 times. Only 2 is excluded. The proof is done by case distinction and can be found in Section 7.2.3.

Now we draw some consequence of this theorem.

Theorem 7.11. *Let E be an elliptic curve $E = E[\mathbb{K}]$.*

1. *$P, Q \in E, P \neq Q$ then there exists a projective line L through P and Q and a 3rd point $R \in E \cap L$.*

2. *$P \in E$, L the tangent line of E at P then there exists another point $R \in E \cap L$.*

In the Statements 1 and 2 the multiplicities have to be taken into account.

"Multiplicities have to be taken into account" means for example in case if the line fixed by P and Q happens to be a tangent line at Q then Q has to be taken with multiplicity 2, hence the additional point R will coincide with Q.

Proof. Case 1: For P and Q with $P \neq Q$ there is a line L passing through P and Q. Hence $\sum m(P', L, E) \geqq 2$. and by Theorem 7.10 the sum has to be $\sum m(P', L, E) = 3$. This shows that there is a 3rd point.
(a) If $m(P, L, E) = m(Q, L, E) = 1$ then the 3rd point R is different from P and Q.
(b) If one of the points has multiplicity 2 with respect to this line, then this point will be equal to the 3rd point R.
(c) Two points with multiplicity equal 2 are not possible as otherwise $\sum m(P', L, E) \geqq 4$ which is a contradiction.

Case 2: Let P be the point and L the (unique) tangent line on E at P. Then $m(P, L, E) \geqq 2$. Hence there is a 3rd point R of intersection. If $m(P, L, E) = 2$ then $R \neq P$ and $m(P, L, E) = 1$. If $m(P, L, E) > 2$ then the 3rd point coincides with P and $m(P, L, E) = 3$. □

Now we are ready to define the group structure.

1. First recall that $O = (0 : 1 : 0)$ is the unique point at infinity of the curve E. Let P be another point on the line at infinity L_∞ i.e. $P = (1 : y : 0)$. We obtain from (7.7)

$$\psi(t) = f(O + t \cdot P) = -t^3,$$

(note that all terms which include Z will vanish), hence $m(O, L_\infty, E) = 3$. In particular, the line at ∞ is the tangent line at the point O at ∞, The multiplicity is 3.

2. Let $P, Q \in E, P \neq Q$. Take the line L_1 through P and Q. Following the theorem above we get a third point R in the intersection $E \cap L_1$. This 3rd point we denote by $P \times Q$. Now we take the line L_2 passing through $P \times Q$ and O. The 3rd point of intersection we take as $P \oplus Q$.

3. For the case $P = Q$, i.e. if we want to define $P \oplus P$ we take as L_1 the tangent at P. Again by the theorem there is another point on the line and on the elliptic curve. This point we denote by $P \times P$. The 3rd point on the line through $P \times P$ and O gives $P \oplus P$.

4. In case that the above constructed lines meet with multiplicities greater than 1 at one of the points then the 3rd point will be this point.

These rules are the rules to define an "addition" on the set of $\mathbb{K}$-valued points of the elliptic curve. A first remark is that the addition is commutative $P \oplus Q = Q \oplus P$, as the line L_1 does not see the order of the points. Second, if one of the points is O we get that on the line L_1 defined by P and O the 3rd point is $P \times O$. Hence our line

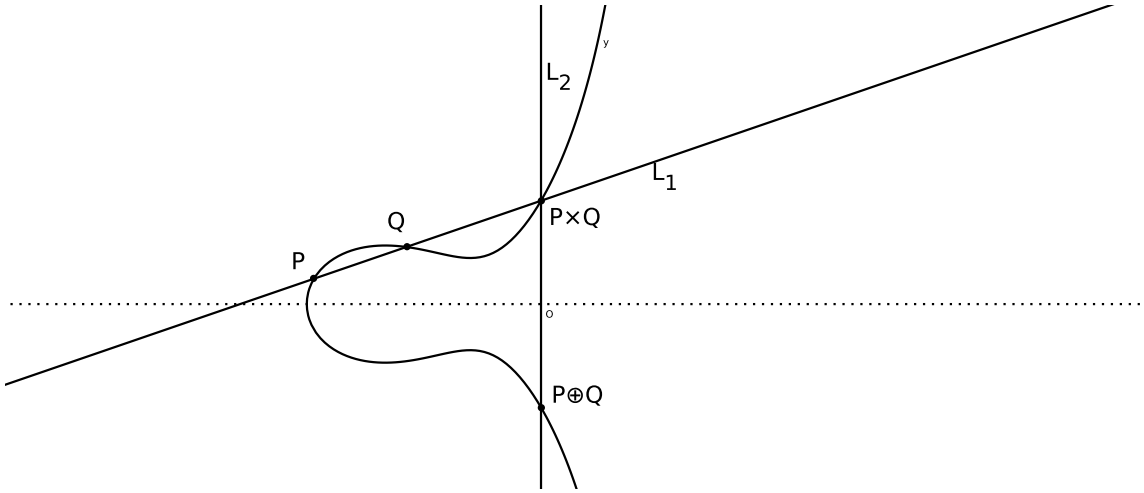

Figure 7.2:

$L_2 = L_1$ and we obtain $P \oplus O = P$ at least if $P \neq O$. To construct $O \oplus O$ we take the tangent at O. As shown above this tangent line is the line at ∞. Furthermore the 3rd point is again O (as the multiplicity is 3). Hence $O \times O = O$ and again L_3 is the line at infinity and hence $O \oplus O = O$.

Theorem 7.12. *The set of $\mathbb{K}$-valued points $E = E(\mathbb{K})$ together with the operation $\oplus$ introduced above gives $(E(\mathbb{K}), \oplus)$ the structure of an abelian group with neutral element O, which is the point at infinity on the curve.*

It remains to show:

1. every element P has an inverse element $\ominus P$, i.e. $P \oplus (\ominus P) = O$.
2. The addition is associative, i.e. $(P \oplus Q) \oplus R = P \oplus (Q \oplus R)$.

To find the inverse element we first show

Lemma 7.13. a) *Assume $P, Q, R \in E$ different. If P, Q, R lie on a line then*

$$(P \oplus Q) \oplus R = O.$$

b) *The statement is also true if they are not necessarily different but then counted with multiplicities, i.e. the line is then tangent at the double point.*

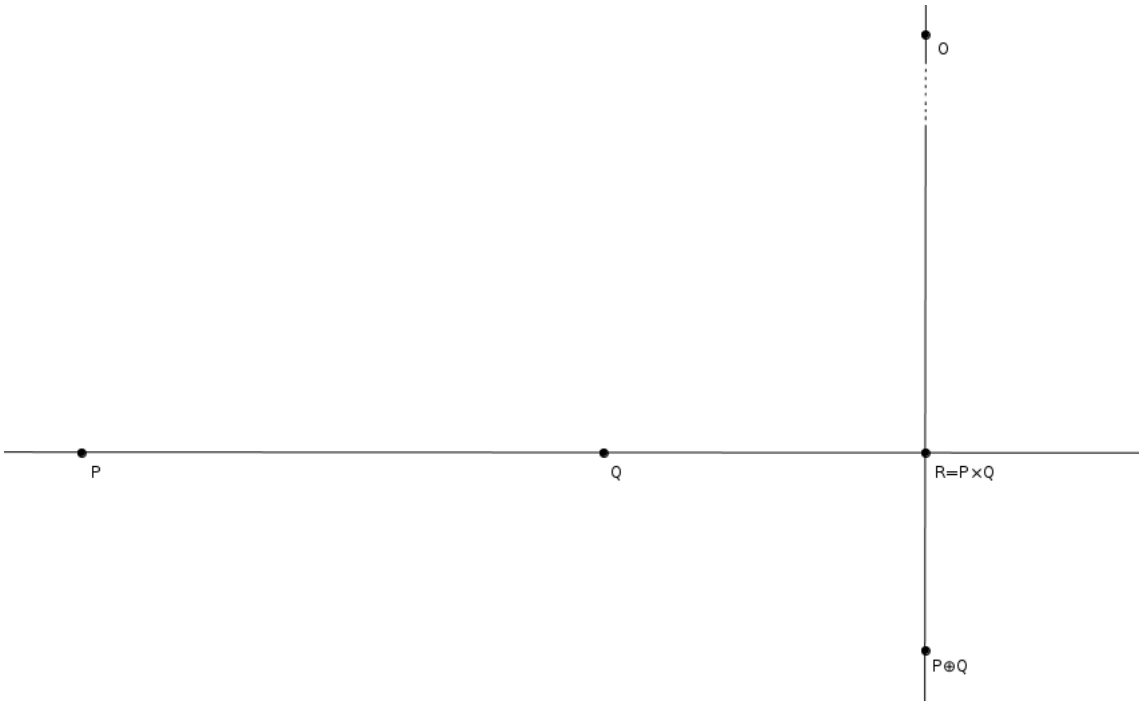

Figure 7.3:

Proof. See Figure 7.3. If P, Q are given and L a line passing through them, then $L = L_1$ is the first line in the construction of $\oplus$. In particular $R = P \times Q$. Next we want to build $(P \oplus Q) \oplus R$. First we have to take the line through the points $(P \oplus Q)$ and R and the 3rd point of intersection yields $(P \oplus Q) \times R$. But this 3rd point is O. The line through O and O is the tangent at O and O is a point of multiplicity 3. Hence the 3. point is also O. We obtain $(P \oplus Q) \oplus R = O$ as required. □

Proposition 7.14. *Let $P \in E(\mathbb{K})$. Take $\ominus P$ to be the 3rd point of intersection of E with the line through P and O then $\ominus P$ is the additive inverse of P, i.e. $P \oplus (\ominus P) = O$. In particular each point $P \in E(\mathbb{K})$ has an inverse element.*

Proof. For the point $\ominus P$ we obtain by the above lemma

$$O = (P \oplus O) \oplus (\ominus P) = P \oplus (\ominus P)$$

as claimed. □

It remains to show associativity: This is the only more complicated part.

Lemma 7.15. *Given an elliptic curve E and 2 pairs of triples of projective lines M_1, M_2, M_3 and L_1, L_2, L_3 such that all intersection points $M_i \cap L_j$ are different (this*

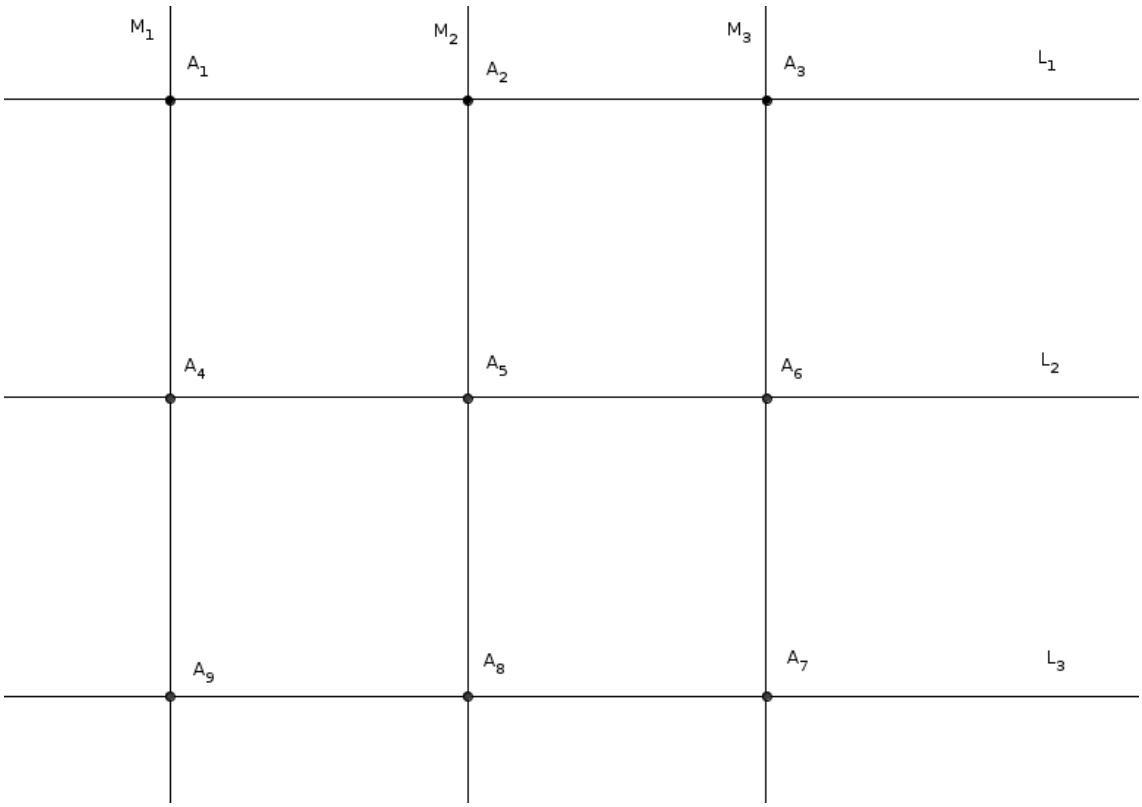

Figure 7.4:

says we have 9 intersection points). Then if 8 of them lie on the elliptic curve E then the point no. 9 will also lie on E .

Proof. First of course Figure 7.4 is misleading as there will be also intersections between the M_i and between the L_j themselves, but in these points we are not interested. We will prove the lemma in Section 7.2.2.

Using this lemma we will show the associativity at least if $O, P, Q, R, P \times Q, Q \times R, P \oplus Q, Q \oplus R$ are pairwise distinct. We refer to Figure 7.5. The points A_1 to A_8 are the points with the labels in the figure. The points A_1 to A_8 are on the corresponding lines and on E by the construction of the points. Hence by the lemma $A_9 = M_1 \cap L_3$ is also on E. By definition of the group structure $P \times (Q \oplus R)$ lies on M_1 and on E, $(P \oplus Q) \times R$ lies on L_3 and E . Now we have a line M_1 and a cubic

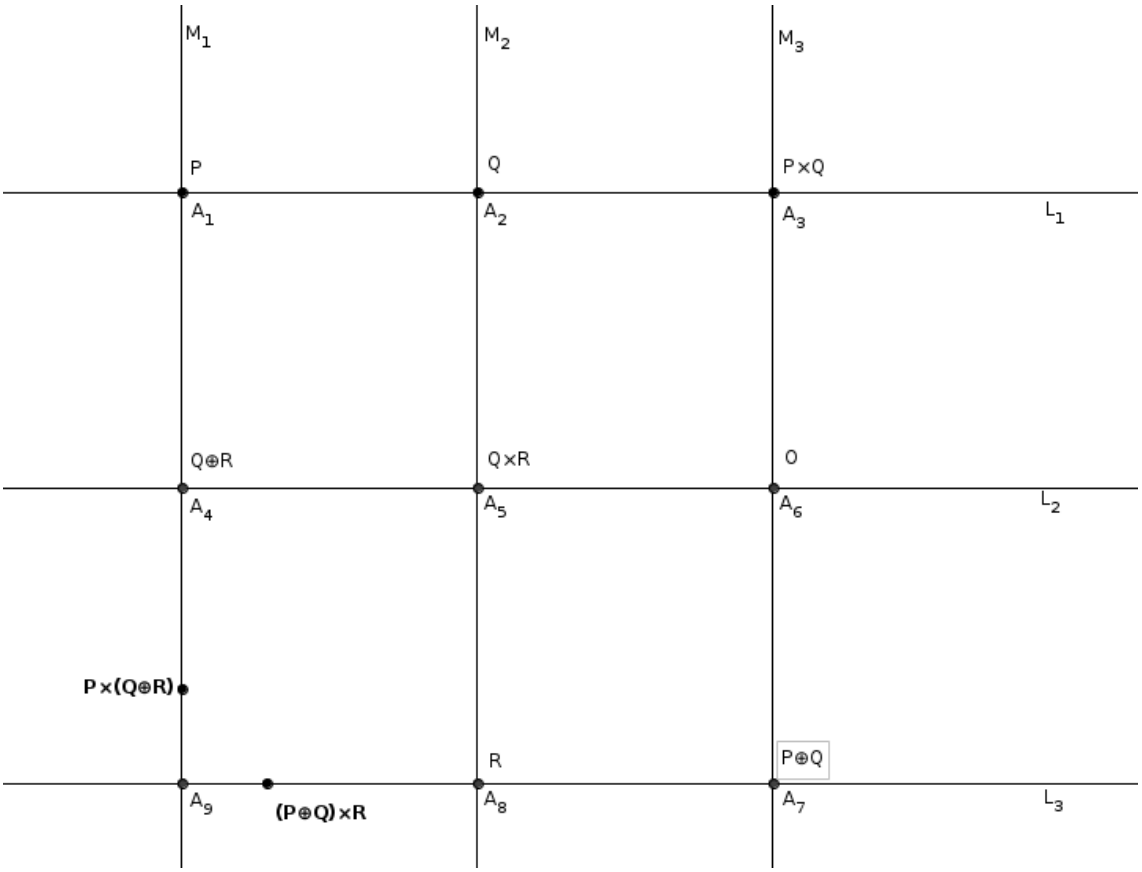

Figure 7.5:

curve E (irreducible). Hence we have maximally 3 points of intersection. One set of intersection points is $P, Q \oplus R, P \times (Q \oplus R)$ another set $P, Q \oplus R, A_9$. Hence the 3rd points have to coincide, i.e.

$$A_9 = P \times (Q \oplus R)\,.$$

In the same way for L_3

$$A_9 = (P \oplus Q) \times R\,.$$

But

$$(P \oplus Q) \oplus R = P \oplus (Q \oplus R) \quad \leftrightarrow \quad (P \oplus Q) \times R = P \times (Q \oplus R)\,.$$

The latter relation is true, hence the first one too. This is the associativity. □

7.2.1 Calculations

The group operations can be done easily in coordinates (with respect to the normal forms (7.1) and (7.2)). First for our neutral element $O = (0 : 1 : 0)$, $-O = (0 : 1 : 0)$ and $O \oplus P = P$. Hence the interesting things only happen in the affine part. Let $P_1 = (x_1 : y_1 : 1)$ then its affine coordinates are $P_1 = (x_1 : y_1)$.

Inverse element

$$-P_1 = (x_1, -y_1 - a_1x_1 - a_3), \tag{7.8}$$

in case $\operatorname{char}\mathbb{K} \neq 2, 3$ we have $a_1 = a_3 = 0$ hence $-P_1 = (x_1, -y_1)$. Why (7.8) is true?

We have to take the line through $P_1 = (x_1 : y_1 : 1)$ and $O = (0 : 1 : 0)$ and the 3rd point will be $-P_1$. The points on the line have the coordinates (x_1, y) with y arbitrary. In particular the x-coordinate is fixed to be x_1. The y coordinate of the inverse has to be calculated from the defining equation

$$y^2 + a_1xy + a_3y - (x^3 + a_2x^2 + a_4x + a_6) = 0$$

for the elliptic curve. If we set

$$c = a_1x_1 + a_3$$

$$d = -x_1^3 - a_2x_1^2 - a_4x_1 - a_6$$

we obtain the quadratic equation

$$y^2 + cy + d = 0.$$

This equation has 2 solutions (maybe with multiplicities) $y_1^{(1)}$ and $y_1^{(2)}$ in the algebraic closure. We know that one exists in $\mathbb{K}$, e.g. $y_1^{(1)}$ as we start from a point on the curve, hence the other also exists in $\mathbb{K}$. Now

$$(y - y_1^{(1)}) \cdot (y - y_1^{(2)}) = y^2 + cy + d$$

and consequently

$$-y_1^{(1)} - y_1^{(2)} = c = a_1x_1 + a_3.$$

This yields $y_1^{(2)} = -y_1^{(1)} - a_1x - a_3$ which was the claim. If $\operatorname{char}\mathbb{K} \neq 2, 3$ we have the nice Figure 7.6.

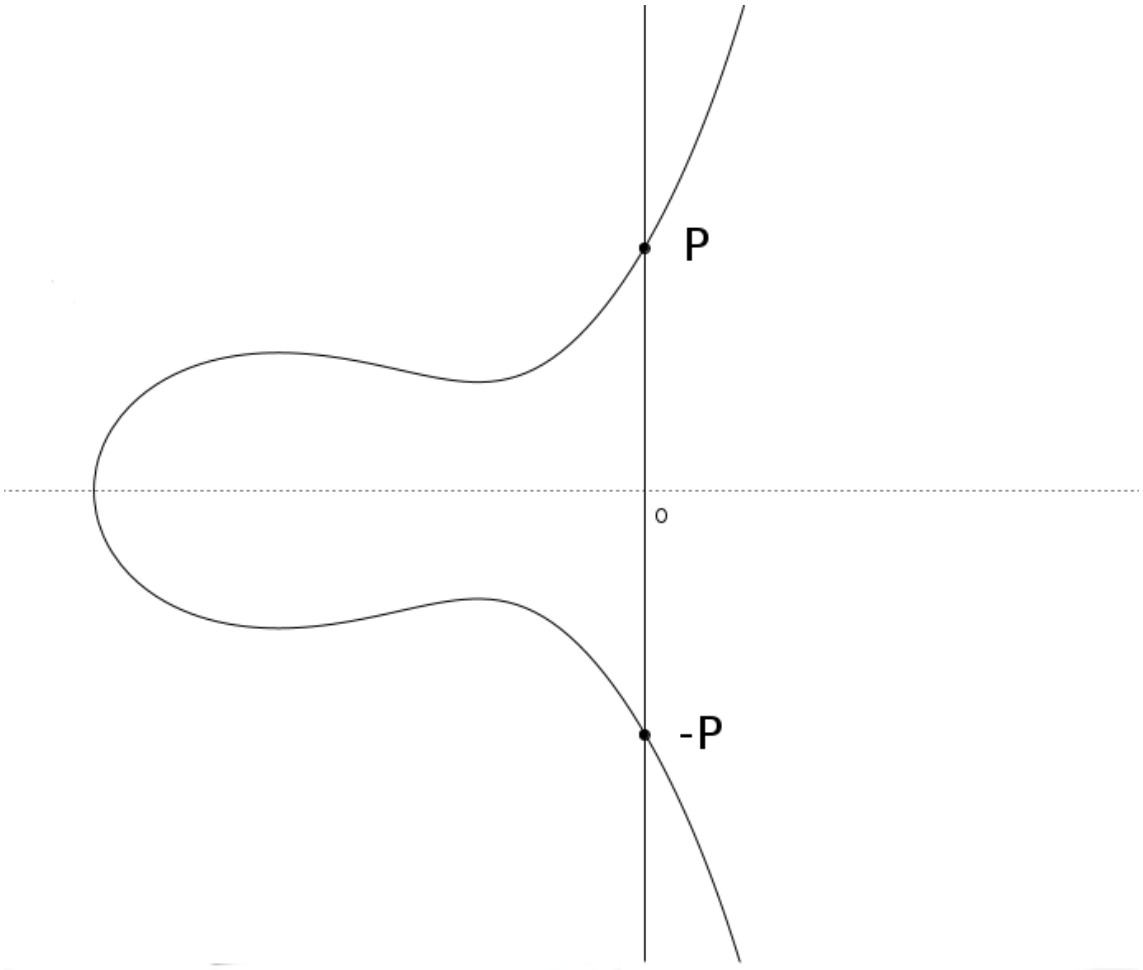

Figure 7.6:

If $P_2 = (x_2, y_2)$, $P_1 = (x_1, y_1)$ then $P_1 + P_2 = P_3 = (x_3, y_3)$ where the x_3 and y_3 can be calculated only using multiplication, additions and divisions. We copy the result from the book of Enge, p.42, Table 2.3, [Enge].

$$x_3 = \begin{cases} \left(\frac{y_2-y_1}{x_2-x_1}\right)^2 + a_1\left(\frac{y_2-y_1}{x_2-x_1}\right) - a_2 - x_1 - x_2 & \text{if} \quad P \neq Q \\ \left(\frac{3x^2+2a_2x+a_4-a_1y}{2y+a_1x+a_3}\right)^2 + a_1\left(\frac{3x^2+2a_2x+a_4-a_1y}{2y+a_1x+a_3}\right) - a_2 - 2x & \text{if} \quad P = Q \end{cases}$$
$$y_3 = \begin{cases} \frac{y_2-y_1}{x_2-x_1}(x_1 - x_3) - y_1 - (a_1x_3 + a_3) & \text{if} \quad P \neq Q \\ \frac{3x^2+2a_2x+a_4-a_1y}{2y+a_1x+a_3}(x - x_3) - y - (a_1x_3 + a_3) & \text{if} \quad P = Q. \end{cases} \tag{7.9}$$

Here we used $x = x_1 = x_2$ and $y = y_1 = y_2$ for $P = Q$.

In case you are interested how to obtain these formulas see [Enge] or [Wer, p.47-53].

If $\operatorname{char}\mathbb{K} \neq 2,3$ we can reach $a_1 = a_2 = a_3 = 0$ and the above expressions simplify to

$$\begin{aligned} x_3 &= \begin{cases} \left(\frac{y_2-y_1}{x_2-x_1}\right)^2 - x_1 - x_2 & \text{if} \quad P \neq Q \\ \left(\frac{3x^2+a_4}{2y}\right)^2 - 2x & \text{if} \quad P = Q \end{cases} \\ y_3 &= \begin{cases} \frac{y_2-y_1}{x_2-x_1}(x_1 - x_3) - y_1 & \text{if} \quad P \neq Q \\ \frac{3x^2+a_4}{2y}(x - x_3) - y & \text{if} \quad P = Q. \end{cases} \end{aligned} \tag{7.10}$$

7.2.2 Proof of Lemma 7.15

We do not repeat the formulations and notation of the lemma. Let V be the vector space of homogeneous polynomials of degree k in n variables, then

$$\dim V = \binom{n-1+k}{k}.$$

Here the degree is 3 and the number of variables is also 3. In particular we obtain $\dim V = 10$ for the space of cubic polynomials. Let V' be the subspace of those cubic polynomials which are zero at all $A_i\,, i = 1, .., 8\,.$

Exercise: verify that is indeed a subspace.

We want to determine the dimension of V'. For this goal we consider the linear maps ψ_s defined as follows (here $s \in \mathbb{P}^2(\mathbb{K})$)

$$\psi_s : V \longrightarrow \mathbb{K}\,; \quad q \longmapsto q(s)\,.$$

The maps ψ_s are the evaluation of the polynomials q at the point s. These are linear maps (in fact linear forms), i.e.

$$\psi(q_1 + q_2) = \psi(q_1) + \psi(q_2) \quad \text{as} \quad (q_1 + q_2)(s) = q_1(s) + q_2(s)\,.$$

For a fixed s we find a cubic polynomial q with $q(s) \neq 0$, hence we have$\psi_s \neq 0\,.$ In particular, ψ_s is not the zero map and hence as a linear form ψ_s is surjective. As

$$\dim\ker\psi_s = \dim V - \dim\operatorname{im}\,\psi_s = 10 - 1 = 9\,.$$

The elements of $\ker \psi_s$ consists of cubic polynomials vanishing at the point s. Hence, by definition

$$V' = \bigcap_{s \in A_1, A_2, .., A_8} \ker \psi_s \, .$$

If we go step by step down each form reduces the previous dimension of the common kernel by at most one. Hence $\dim V' \geq 2$.

Claim $\dim V' = 2$.

Proof. Let $S \in L_1 \cap L_2$. As two projective lines always have an intersection point such a point S exists. Moreover, $S \notin \{A_1, A_2, .., A_8\}$. Let $V_S = \ker \psi_S$. If we can

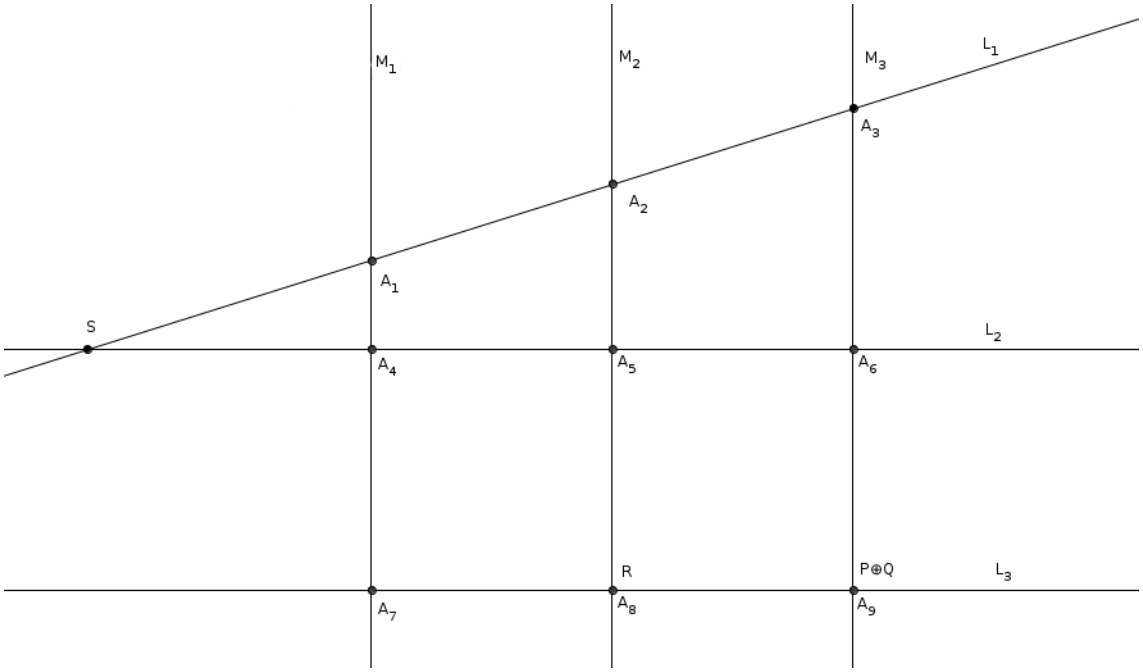

Figure 7.7:

show that $\dim(V' \cap V_S) = 1$ then $\dim V' = 2$ as the dimension by cutting with $\ker \psi_S$ can maximally drop by 1.

Let p be an element of $V' \cap V_S$ (i.e. a cubic polynomial) then $p(A_1) = p(A_2) = p(A_3) = p(S) = 0$ and $A_1, A_2, A_3, S \in L_1$. Hence $\mathfrak{V}(p) \cap L_1$ has at least 4 intersection points. If L_1 is not a component of $\mathfrak{V}(p)$ then the maximal number of intersection points is 3. Hence L_1 is a component of $\mathfrak{V}(p)$. The same is true for L_2. Being a

component corresponds to writing the polynomial p as factor $p = l_1 \cdot l_2 \cdot l$ with $L_1 = \mathfrak{V}(l_1)$, $L_2 = \mathfrak{V}(l_2)$ and $L = \mathfrak{V}(l)$ another line. But $p \in V'$ hence $p(A_7) = p(A_8) = 0$. As $A_7, A_8 \notin L_1 \cup L_2$ they have to be zeros of the complimenting factor which is given by L_3, hence $L = L_3$, i.e. $p = l_1 \cdot l_2 \cdot l_3$ up to multiplication with a scalar. Hence every such p will be a multiple of $l_1 \cdot l_2 \cdot l_3$. Consequently, $\dim(V_S \cap V') = 1$ and the claim follows. □

Now we give the proof of the Lemma 7.15.

Proof. Let the lines M_i be given by the linear polynomials m_i and the lines L_j by l_j . We consider the two cubic polynomials

$$p_1 = m_1 \cdot m_2 \cdot m_3 \qquad p_2 = l_1 \cdot l_2 \cdot l_3 \,.$$

As they cannot be multiples of each others they are linearly independent. By construction

$$p_1(A_i) = p_2(A_i) = 0 \,, \quad i = 1, .., 9 \,.$$

In particular this implies that p_1 and $p_2 \in V'$ (we take only $A_1, \ldots, A_8$ into account). And as $\dim V' = 2$ they constitute a basis of V' . Our defining polynomial g for the elliptic curve vanishes also for the 8 points $A_1, .., A_8$, hence $g \in V'$. This means that g is a linear combination of p_1 and p_2 , i.e.

$$g = \alpha p_1 + \beta p_2 \,, \quad \alpha, \beta \in \mathbb{K}.$$

Hence

$$g(A_9) = \alpha p_1(A_9) + \beta p_2(A_9) = 0 + 0 = 0 \,.$$

This was the claim. □

7.2.3 Proof of Theorem 7.10

For the convenience of the reader we repeat the formulation of the theorem

Theorem. *Let L be a projective line and E an elliptic curve then*

$$\sum_{P \in E \cap L} m(P, L, E) = \begin{cases} 0 \\ 1 \\ 3 \end{cases} .$$

Note that the sum over the points of the plane reduces to the sum over the points in $E \cap L$. The defining polynomials are

$$\begin{aligned} L: &\quad l(X,Y,Z) = aX + bY + cZ, \\ E: &\quad g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3. \end{aligned}$$

As we already made a transformation of variables to obtain the normal form for E we cannot adjust the defining polynomial for L any further.

We have to make case distinction depending on the line L.

Case 1: (in the defining polynomial for L we have $a = 0 = b$). This says that $L = \mathfrak{V}(Z)$. In other words, L is the line at infinity. Hence

$$L \cap E = \{(\alpha : \beta : 0) \mid g(P) = 0\}. \tag{7.11}$$

If we plug this into the defining polynomial g for E we obtain

$$g(\alpha : \beta : 0) = -\alpha^3 .$$

Hence the only point is $(\alpha : \beta : \gamma) \in L \cap E = (0 : 1 : 0) = O$. To check the multiplicity we have to choose another point P' on L (the choice is arbitrary). We take $P' = (1 : 0 : 0)$; clearly $l(P') = 0$. Now

$$\psi(t) = g\left((0,1,0) + t((1,0,0)\right) = g(t,1,0) = -t^3.$$

Hence, the point $(0 : 1 : 0)$ has multiplicity 3, and the claim is true for this special line L.

Case 2: (in the defining polynomial for L we have $a \neq 0$, $b = 0$). Hence, L is given by $l(X,Y,Z) = a \cdot X + c \cdot Z$ (here $c = 0$ is not excluded). First note that the point $O = (0 : 1 : 0) \in L$. Now let $P \in L$, with $P = (\alpha : \beta : \gamma)$. As $P \in L$ we have $a \cdot \alpha + c \cdot \gamma = 0$, (i.e. $\alpha = -\frac{c}{a} \cdot \gamma$) which says $P = (-\frac{c}{a} \cdot \gamma : \beta : \gamma)$. Now we make case distinctions with respect to the point P.

Case 2a: ($\gamma = 0$). This implies $P = (0 : 1 : 0) = O$ and P lies on E. For calculating its multiplicity with respect to the line L we take e.g. the point $P' = (-c : 0 : a) \in L$ and obtain

$$\psi(t) = g(O + t \cdot P') = g(-ct, 1, at) = a \cdot t + O(t^2) .$$

Hence $m(P, L, E) = 1$

Case 2b: ($\gamma \neq 0$). Of course we require that P has to lie on E. In particular also $P \neq O$. It might be that there are no such points. In case that we have such a point we can normalize it to be $P = (-\frac{c}{a} \cdot \gamma : \beta : 1)$ where the new β is $\frac{\beta_{\text{old}}}{\gamma}$). We consider the polynomial

$$h(s) = g(-\frac{c}{a}, s, 1),$$

in the variable s. As $P \in E$, i.e. $g(P) = 0$ we get that β is a zero of the polynomial h. Now we take as auxiliary point the point $O = (0 : 1 : 0)$ and obtain

$$\psi(t) = g(P + t \cdot O) = g(-\frac{c}{a}, \beta + t, 1) = h(\beta + t)\,.$$

We write

$$h(s) = (s - \beta)^k \cdot h^*(s)$$

where k is the order of the zero β of the polynomial $h(s)$ and $h^*(\beta) \neq 0$. We write

$$\psi(t) = h(\beta + t) = (\beta + t - \beta)^k \cdot h^*(\beta + t) = t^k \cdot h^*(\beta + t).$$

As $h^*(\beta) \neq 0$ the term $h^*(t + \beta)$ has order zero in t and hence the order of the zero of t in ψ is equal to order of the zero of β in h. Now we check in more detail the polynomial

$$h(s) = g(-\frac{c}{a}, s, 1)\,.$$

The variable s corresponds to the variable Y in the defining equation g for E. As there it is of degree two, $h(s)$ is also a degree two polynomial. Hence, we will either have no zeros, one zero with multiplicity two or two zeros with multiplicity one.

In total we will have from Case (2a) a contribution of 1 to the sum and from Case (2b) either no contribution or a contribution of 2. Hence, also in this case for the line L the relation $\#(E \cap L) = 1$, or 3. Which was to show.

Case 3: (in the defining polynomial for L we have $b \neq 0$). This is the remaining case. In this case the point $O = (0 : 1 : 0)$ does not lie on L. Hence all points of $L \cap E$ lie in the affine part. The point on the line can be written as

$$P = (\alpha : \beta : 1) \in L \quad \leftrightarrow \quad \beta = -\frac{c}{b} - \frac{a}{b} \cdot \alpha\,.$$

These can be parameterized by

$$P_\alpha = (\alpha : -\frac{c}{b} - \frac{a}{b} \cdot \alpha : 1)$$

For $\alpha \in \mathbb{K}$ we get all points on the affine part of the line.

We set $h(\alpha) = g(\alpha, -\frac{c}{b} - \frac{a}{b} \cdot \alpha, 1)$ and obtain that $P_\alpha \in E$ if and only if $h(\alpha) = 0$, Equivalently, α is a zero of $h(s) = g(s, -\frac{c}{b} - \frac{a}{b}s, 1)$. As auxiliary point we take

$$P' = (-b : a : 0) \in L$$

and get

$$\psi(t) = g(P_\alpha + tP') = g(\alpha - tb, -\frac{c}{b} - \frac{a}{b}(\alpha - tb), 1) = h(\alpha - tb)\,.$$

As in the previous case we obtain "multiplicity of t in ψ"= "zero order of α in h". Now $h(s)$ is a polynomial of order 3 in s. (To see this just plug in $(s, -\frac{c}{b} - \frac{a}{b} \cdot \alpha, 1)$ into the defining equation of E given by g and check the order of s.) Such a polynomial has in $\overline{\mathbb{K}}$ (the algebraic closure of $\mathbb{K}$) exactly 3 roots counted with multiplicities

$$h(s) = -(s - \alpha_1)(s - \alpha_2)(s - \alpha_3)\,.$$

We rewrite this to

$$h(s) = -s^3 + (\alpha_1 + \alpha_2 + \alpha_3)s^2 + ..$$

which is now a polynomial with coefficients from $\mathbb{K}$.

How many zero with multiplicities are possible over $\mathbb{K}$? We have three cases: no zero at all, 1 zero or 3 zeros. The possibility of two zeros is excluded, as if $\alpha_1, \alpha_2 \in \mathbb{K}$, then from $(\alpha_1 + \alpha_2 + \alpha_3) = \eta \in \mathbb{K}$ it follows that the 3rd root $\alpha_3 = \eta - (\alpha_1 + \alpha_2)$ is also in $\mathbb{K}$.

Altogether this shows Theorem 7.10. □

7.3 An application of the group structure

Given an elliptic curve E defined over a finite field $\mathbb{F}_q$, $q = p^n$, $p \in \mathbf{P}$ we showed above that on E we have the structure of an abelian group $(E, \oplus)$. The elements of E define a subset of the projective plane over $\mathbb{F}_q$. The number of points in $\mathbb{P}^2(\mathbb{F}_q)$ is finite. In fact, it is equal to $q^2 + q + 1$ [2]. In particular, $(E, \oplus)$ is a finite group. In general it will not be a cyclic group. But if we fix a point $Q \in E$ we obtain the finite cyclic subgroup $\langle Q \rangle$ generated by Q. In fact with a suitable Q we obtain cyclic groups which have many elements. Calculations in E (and hence in $\langle Q \rangle$) are very easy (see the expressions (7.9), (7.10)). Nevertheless, the discrete logarithm

[2] Exercise: show this formula for the number of points in $\mathbb{P}^2(\mathbb{F}_q)$.

problem (see Section 9.1) is very hard to solve - at least if some special curves are excluded. Such groups are needed in cryptography. The method is now widely used.

As an example used in the context of bitcoins (see [Bit]) we quote the following. One takes F_p with p the prime number

$$p = 2^{256} - 2^{32} - 2^9 - 2^8 - 2^7 - 2^6 - 2^4 - 1, \tag{7.12}$$

which is slightly less than 2^{256}. One considers the elliptic curve E (in affine description)

$$Y^2 = X^3 + 7. \tag{7.13}$$

As generator for the cyclic subgroup considered one takes the (affine) point $Q = (q_x, q_y) \in E$ with affine coordinates[3]

$$q_x = 55066263022277343669578718895168534326250603453777594175500187360389116729240,$$
$$q_y = 32670510020758816978083085130507043184471273380659243275938904335757337482424.$$

For more details on the application of elliptic curves in crypto, see [Wer] and [Enge].

[3]Up to copy-errors.

Chapter 8

Elliptic curves over $\mathbb{C}$ and Tori

In this chapter we will discuss elliptic curves over the field of complex numbers $\mathbb{C}$ from an analytic point of view. It will turn out that complex one-dimensional tori can be identified with elliptic curves (over $\mathbb{C}$) as introduced before. We will occasionally use some language and results from complex analysis. Nevertheless, it should be understandable with only limited background. In case that you want to know more, please consult [RS] where the construction is done in detail.

Let ω_1, ω_2 be two complex numbers which are linearly independent over the real numbers. In particular $\lambda_1, \lambda_2 \in \mathbb{R}$ with $\lambda_1\omega_1 + \lambda_2\omega_2 = 0$ implies that $\lambda_1 = \lambda_2 = 0$.

The set

$$\Gamma := \{n\omega_1 + m\omega_2 \mid n, m \in \mathbb{Z}\}$$

is a 2-dimensional lattice in $\mathbb{R}^2 \cong \mathbb{C}$. In particular, Γ is a subgroup of $(\mathbb{C}, +)$, the additive group of complex numbers. We consider the quotient group

$$T = \mathbb{C}/\Gamma.$$

Recall that the quotient group consists of equivalence classes $\bar{z}$ with respect to the equivalence relation

$$z' \sim z \quad \text{if and only if} \quad z' = z + \omega \quad \text{with } \omega \in \Gamma.$$

Recall the definition of the equivalence class as

$$\bar{z} := \{z' \in \mathbb{C} \mid z' \sim z\}.$$

By dividing each complex numbers by ω_1 we can rescale the whole situation without changing the principle. The rescaled lattice will now be generated by the two

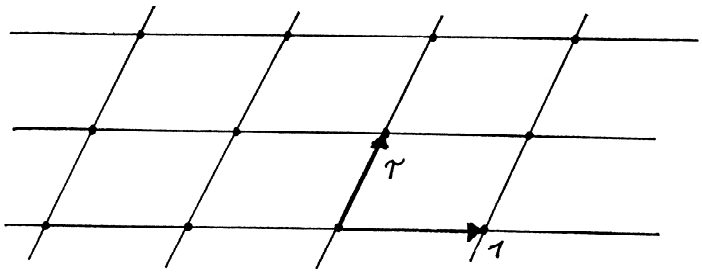

Figure 8.1: A two-dimensional standard lattice

complex numbers

$$1 = \frac{\omega_1}{\omega_1}, \qquad \tau := \frac{\omega_2}{\omega_1}.$$

We will even choose τ such that its imaginary part is strictly positive (we choose as generator instead of ω_2 the element $-\omega_2$ if needed). Such a lattice is called *standard lattice*. For the following we will always assume that the lattice is given as a standard lattice

$$\Gamma := \{n + m\tau \mid n, m \in \mathbb{Z}\}, \tag{8.1}$$

see the corresponding picture Fig. 8.1. The points of the subset

$$\mathcal{F} := \{z \in \mathbb{C} \mid z = a + b\,\tau, \quad 0 \leq a, b < 1, a, b \in \mathbb{R}\}, \tag{8.2}$$

correspond 1:1 to the points of T, i.e. to the equivalence classes. This $\mathcal{F}$ is called fundamental region of T.

The space $\mathbb{C}$ carries a topological and complex structure. We can do complex analysis. We define what holomorphic functions are, what meromorphic functions are, make power (or Laurent) series expansions, and many more things. This complex structure will be exported to the quotient $T = \mathbb{C}/\Gamma$. The T will be a complex manifold of dimension 1, i.e. a Riemann surface. It will be a *complex one-dimensional torus*. Topologically it will be obtained by identifying the two complimentary edges of the fundamental region, see Fig. 8.2. We note the fact, that the complex structure will depend on the value of τ, as the lattice will depend on it.

A function f on the complex plane will be a function on the torus T if and only if f is doubly periodic with respect to the lattice. Doubly periodic means that

$$f(z + n + m\tau) = f(z), \quad \forall n, m \in \mathbb{Z} \qquad \forall z \in \mathbb{C}.$$

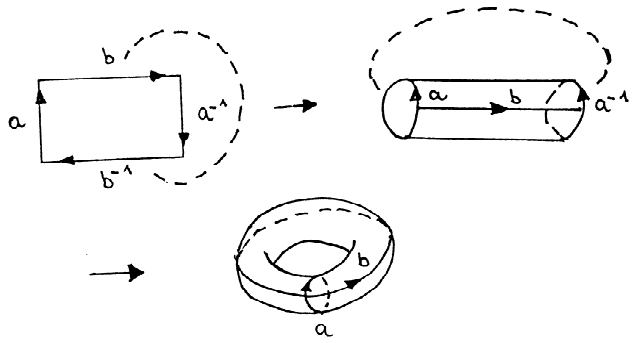

Figure 8.2: Glueing of the torus.

Such f defines a function $\bar{f}$ on T by setting $\bar{f}(\bar{z}) := f(z)$. Vice versa, given g a function on T, we get by defining $f(z) := g(z+\Gamma)$ a function f on $\mathbb{C}$ which is doubly periodic and satisfies $\bar{f} = g$.

This correspondence remains true if we consider holomorphic or even meromorphic doubly periodic functions. Recall that holomorphic functions on $\mathbb{C}$ are functions which have at every point of $\mathbb{C}$ a power series expansion in the complex coordinate z. A meromorphic functions f is defined as holomorphic function on an open dense subset of $\mathbb{C}$ with discrete complement and such that f has maybe algebraic poles of the points of the complement (saying that for those points we have Laurent expansions which have only finitely many negative terms).

As holomorphy and meromorphy are local conditions they transfer directly from $\mathbb{C}$ to T. In particular, the doubly periodic (with respect to Γ) meromorphic functions on $\mathbb{C}$ coincide with the meromorphic functions on the torus $T = \mathbb{C}/\Gamma$. These functions are also called *elliptic functions*. We denote the set of meromorphic functions on T by $\mathcal{M}(T)$. In fact, it is a (transcendental) field extension of the complex field $\mathbb{C}$, as the sum, resp. product of two meromorphic functions is again a meromorphic function. Also for $f \neq 0$ (the zero function) $1/f$ is also a meromorphic function, as f can only have a discrete set of zeros of finite order.

Next we want to study this in detail.

Proposition 8.1. *The only global holomorphic functions on the torus $T = \mathbb{C}/\Gamma$ are*

the constants.

Proof. (Only for those who know a little bit more about complex analysis.) A holomorphic function will be bounded on the fundamental region $\mathcal{F}$. By the double periodicity it will be bounded in $\mathbb{C}$. Hence, by Liouville theorem it will be a constant. □

Proposition 8.2. *There is no meromorphic function on the torus which has exactly one pole of order 1.*

Proof. (Only for those who know a little bit more about complex analysis.) Let $\bar{f}$ be a meromorphic function on the torus, f be the corresponding doubly periodic meromorphic function on $\mathbb{C}$, and $\mathcal{F}$ the fundamental region as defined above (respectively its closure). To start with let us assume that f does not have poles on the boundary $\partial\mathcal{F}$. The classical residue theorem of complex analysis says

$$\frac{1}{2\pi\mathrm{i}} \int\limits_{\partial\mathcal{F}} f(z)\, dz = \sum_{a\in\mathcal{F}} \mathrm{res}_a(f).$$

Due to the double periodicity of f we obtain that the integrals over parallel edges cancel; hence in total

$$\int\limits_{\partial\mathcal{F}} f(z)\, dz = 0,$$

and consequently

$$\sum_{a\in\mathcal{F}} \mathrm{res}_a(f) = 0.$$

A function which has only one pole of order one inside of $\mathcal{F}$ would have a residue. Hence, it cannot exist. In case f has a pole on the boundary $\partial\mathcal{F}$ we deform the region of integration by adding an small open neighbourhood around this special point on one side of $\mathcal{F}$ and subtracting it on the other side and argue as above. □

To describe all functions on the torus we use the *Weierstraß $\wp$−function* (on $\mathbb{C}$)

$$\wp(z) := \frac{1}{z^2} + \sum_{\omega\in\Gamma}{}' \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right).$$

The $'$ denotes that we leave out $\omega = 0$ from the summation. This series converges for all $z \notin \Gamma$ to a holomorphic function. It has poles of second order at the lattice

points. It is doubly periodic and an even function. [1] If we differentiate $\wp$ we obtain

$$\wp'(z) = -\sum_{\omega\in\Gamma} \frac{2}{(z-\omega)^3}.$$

The function $\wp'$ is obviously doubly periodic and has poles of order 3 at the lattice points and is an odd function.

Of course $\wp$ and $\wp'$ are linearly independent, but they are not algebraically independent.

If we compare the coefficients of the power series for $\wp$ and $\wp'$ we get the relation

$$(\wp')^2 = 4\wp^3 - g_2\wp - g_3, \tag{8.3}$$

with

$$g_2 = 60 \sum_{\omega\in\Gamma}{}' \frac{1}{\omega^4} = \sum_{n,m\in\mathbb{Z}}{}' \frac{1}{(n+m\tau)^4},$$
$$g_3 = 140 \sum_{\omega\in\Gamma}{}' \frac{1}{\omega^6} = \sum_{n,m\in\mathbb{Z}}{}' \frac{1}{(n+m\tau)^6}.$$

These g_2, g_3 are called *Eisenstein series*. If the lattice is fixed then they are constants. But they vary with the lattice. For our standard lattices they are holomorphic functions in the variable τ (for $\operatorname{im}\tau > 0$).

Starting with the Eisenstein series we are able to construct the *(analytic) discriminant function*

$$\widetilde{\Delta}(\tau) = {g_2}^3(\tau) - 27{g_3}^2(\tau).$$

An important result is that $\widetilde{\Delta}(\tau) \neq 0$, and we define

$$j(\tau) = 1728 \cdot \frac{{g_2}^3(\tau)}{\widetilde{\Delta}(\tau)}.$$

This j function classifies the isomorphy classes of tori. It is also called the *elliptic modular function*.

Theorem 8.3. *The field of meromorphic functions on the torus, or equivalently the field of doubly periodic (i.e. elliptic functions) can be given as*

$$\mathcal{M}(T) = \mathbb{C}(\wp, \wp')$$

with the relation (8.3). In a more algebraic notation this writes as

$$\mathcal{M}(T) \cong \mathbb{C}(\mathbf{X})[\mathbf{Y}]/(\mathbf{Y}^2 - 4\mathbf{X}^3 + g_2\mathbf{X} + g_3).$$

[1] see Hurwitz-Courant,[HC], p.161 for these results.

Remarks on the notation:
(1) $\mathbb{C}(\wp, \wp')$ denotes the field of quotients of polynomial expressions in $\wp$ and $\wp'$.
(2) $\mathbb{C}(X)$ denotes the field of rational functions in the formal variable X, i.e. the field of quotients of polynomials, or equivalently $\mathbb{C}(X) = \operatorname{Quot} \mathbb{C}[X]$.

Proof. Let $f \in \mathcal{M}(T)$ be given with a pole of order m at the point $\bar{a} \neq \bar{0}$.

$$g(\bar{z}) = f(\bar{z}) \cdot (\wp(\bar{z}) - \wp(\bar{a}))^m$$

is now another function which has this pole removed. By induction we reach a function having only poles at $\bar{0}$. Subtracting complex multiples of $\wp$ and $\wp'$ and taking into account that there are no doubly periodic functions which have only a pole of order one (see Proposition 8.2, we get finally an everywhere holomorphic function on the torus, hence a constant. Working backwards we get the claim. □

Next we want to identify T with an elliptic curve in $\mathbb{P}^2(\mathbb{C})$. We consider the cubic curve $E = \mathfrak{V}(f)$ with

$$f(X, Y, Z) = Y^2Z - 4X^3 + g_2XZ^2 + g_3Z^3, \quad \text{with } \widetilde{\Delta} = g_2^3 - 27g_3^2 \neq 0. \tag{8.4}$$

To make contact with the Weierstraß normal forms of the kind which we discussed in the previous chapter we rewrite the polynomial $-f/4$ (which gives of course the same E) as

$$-\left(\frac{Y}{2}\right)^2 Z + \left(X^3 - \frac{g_2}{4}XZ^2 - \frac{g_3}{4}Z^3\right).$$

If we set $Y' = Y/2$, $a_4 = -g_2/4$, and $a_6 = -g_3/4$ we obtain our previous form (with variables Y' and X) (7.2). The discriminant (7.3) calculates as

$$\Delta = 4a_4^3 + 27a_6^2 = -\frac{1}{4^2}(g_2^3 - 27g_3^2) = -\frac{1}{4^2}\widetilde{\Delta}.$$

As $\widetilde{\Delta} \neq 0$ also $\Delta \neq 0$ and our cubic curve (8.4) is non-singular. Which says that it is an elliptic curve.

We define the following map

$$\Psi : T \to \mathbb{P}^2, \qquad \bar{z} \mapsto \begin{cases} (\wp(z) : \wp'(z) : 1), & \bar{z} \neq \bar{0}\,, \\ (0 : 1 : 0), & \bar{z} = \bar{0}\,, \end{cases}$$

where $\bar{z} = z + L$. Recall that $\bar{0}$ is the class of lattice points. Furthermore note that $\wp(z+L) = \wp(z)$. By the differential equation of the Weierstraß $\wp$ function (8.3) the images of the non-lattice points lie on the affine part of the elliptic curve E (8.4). The class of the lattice points maps to the point at ∞ of E.

Proposition 8.4. *The map Ψ is an analytic isomorphism of T with E.*

Proof. $\wp$ is an analytic function on the torus. As such it can be considered as a map to $\mathbb{P}^1(\mathbb{C})$; the poles are mapped to $\infty \in \mathbb{P}^1(\mathbb{C})$. From the theory of analytic functions on Riemann surfaces one knows that it takes every value of $\mathbb{P}^1(\mathbb{C})$ equally often (calculated with multiplicity). It has a pole of order 2 at $\bar{0} \in T$ and nowhere else. Hence every value occurs two times. But $\wp$ is an even function hence the two points with the same value are $\bar{z}$ and $-\bar{z}$.
Injectivity. Assume that $(\wp(z), \wp'(z)) = (\wp(w), \wp'(z))$ then $\bar{w} = -\bar{z}$ and hence $\wp'(w) = \wp'(-z) = -\wp'(z)$. If $\wp'(z) \neq 0$ the second components cannot be the same. It remains to study the the zeros of $\wp'$. As $\wp'$ has poles of order 3 at the lattice points (hence at $\bar{0} \in T$) it has also 3 zeros on T. These zeros are $1/2,\ \tau/2,\ (1+\tau)/2$. Why? We calculate (using the fact that $\wp'$ is doubly-periodic)

$$\wp'(1/2) = \wp'(1/2 - 1) = \wp'(-1/2) = -\wp(1/2),$$

hence $-\wp(1/2) = 0$. Similar calculations work for the other two points. For these three points we have $\bar{z} = -\bar{z}$. Hence, there is only one point of the torus which maps to the point on the curve we started with. This shows Injectivity on the affine part. Injectivity at the point $\bar{0}$ is by definition of the map.
Analyticity. On the affine part the map is given by the holomorphic functions $\wp$ and $\wp'$. Hence the statement is clear. For checking it around 0 we rewrite

$$\Psi(\bar{z}) = (\frac{\wp(z)}{\wp'(z)} : 1 : \frac{1}{\wp'(z)})$$

in a neighbourhood of z not containing the zeros of $\wp'$. This expression also makes sense for $z = 0$ in the limit and yields $(0 : 1 : 0)$. Hence, Ψ is also analytic at 0, and $\bar{0}$ respectively.
Surjectivity. Let us take (α, β) a point in the affine part of the elliptic curve E. In particular, the coefficients fulfill the equation of the curve. The doubly periodic function $\wp(z) - \alpha$ has a pole of order two at the lattice points, hence it has also two zeros in $\mathcal{F}$ (respectively on T). Let z_1 be such a zero. If we plug in this point

into the differential equation for $\wp$ we get that $\wp'(z_1)^2 = \beta^2$. Hence, $\beta = \pm\wp'(z_1)$. If $\beta = +\wp(z_1)$ we are done, as $\Psi(z_1) = (\wp(z_1), \wp'(z_1)) = (\alpha, \beta)$. Otherwise, we take as pre-image $z_2 = -z_1$ and obtain $\Psi(z_2) = (\wp(-z_1), \wp'(-z_1)) = (\wp(z_1, -\wp'(-z_1)) = (\alpha, \beta)$. As our curve has only one point at ∞, the point $(0 : 1 : 0)$ which is the image point of $\bar{0}$ we showed surjectivity. □

Remark 8.5.

1. In this way we showed that each complex 1-dimensional torus is an elliptic curve over $\mathbb{C}$. Indeed it is also possible to show the opposite. In particular, from the complex analytic point of view they are the same objects.

2. Note that ψ is not an algebraic isomorphism as it is given by transcendental functions.

3. The identification gives another interpretation of the group law on elliptic curves. Note that T has a group structure coming from the group

$$(\mathbb{C}, +): \qquad \bar{z}_1 + \bar{z}_2 = \overline{z_1 + z_2}\,.$$

If we transfer this group law via Ψ we get a group law on the elliptic curve. In fact, it is exactly the group law we have expressed in (7.10). To show that they coincide one needs some additional arguments which we will not reproduce here. As a consequence from Ψ we obtain the so-called addition theorem for the $\wp$ function[2]. We obtain for $z_1 \neq z_2$

$$\wp(z_1 + z_2) = \frac{1}{4}\left(\frac{\wp'(z_1) - \wp'(z_2)}{\wp(z_1) - \wp(z_2)}\right)^2 - (\wp(z_1) + \wp(z_2))\,. \tag{8.5}$$

$$\wp(2z) = -2\wp(z) + \frac{1}{4}\left(\frac{6\wp^2(z) - 1/2\; g_2}{\wp'(z)}\right)^2. \tag{8.6}$$

Recall that the x-coordinate of the point on the elliptic curve is given by the $\wp$ function. If we write down the corresponding formula (7.10) for the y-coordinate we get similar expressions for an addition theorem for $\wp'$.

[2]The factor of $1/4$ in the first expression comes from the fact that our $Y' = \wp'/2$.

Chapter 9

Mixed topics

9.1 The discrete logarithm problem (DLP)

Let G be a finite group with neutral element e. Let a be an element of G and $\langle a \rangle$ the cyclic subgroup of G generated by a, i.e.

$$\langle a \rangle \ := \{a^n \,|\, n \in \mathbb{Z}\}\,.$$

Of course $\langle a \rangle$ will only have finitely many different elements. Set $\mathrm{ord}(a) = \#\langle a \rangle$. Then

$$\langle a \rangle \ := \{a^0 = e, a^1, a^2, .., a^{\mathrm{ord}(a)-1}\}\,.$$

The group G itself is called a *cyclic group* if an $a \in G$ exists with $G = \langle a \rangle$. Such an a is called a *generator* of G. Given $a \in G$ and $k \in \mathbb{N}$ then calculating $b = a^k$ is simple (assuming that we have a group in which we can easily do multiplication).

But: Given b and a searching k such that $a^k = b$ is typically difficult. Finding such a k is called the *discrete logarithm problem (DLP)* as "$k = \log_a b$" . One "brute force" approach is to calculate all a^k for all k and compare the result with b. This is not feasible if our group G is large.

Before discussing suitable groups I like to show how this can be used to exchange a secret key c between partners A and B. The procedure is called *Diffie-Helmann key exchange*.

1.The parties A and B choose a cyclic group G and a generator a, i.e. the pair (G, a). These data can be revealed to everybody.

2. Now A chooses a secret key k and calculates a^k, and B chooses a secret key l and calculate a^l.

3. They exchange the calculated values.

4. Hence A receives a^l, takes its kth power and calculates $(a^l)^k = a^{l \cdot k}$. The party B receives a^k, takes its lth power and calculates $(a^k)^l = a^{l \cdot k}$.

5. The common value $c = a^{l \cdot k}$ is their shared secret.

6. Known (as it went over the transmission channel) are a, a^l, a^k but neither k, nor l nor $c = a^{l \cdot k}$. To determine c one would have to know l or k. In other words, one would need to calculate the discrete logarithm of a^l or a^k.

Encoding (this means calculating a^k) can be done very effectively by squaring and multiplying as the Figure 9.1 shows.

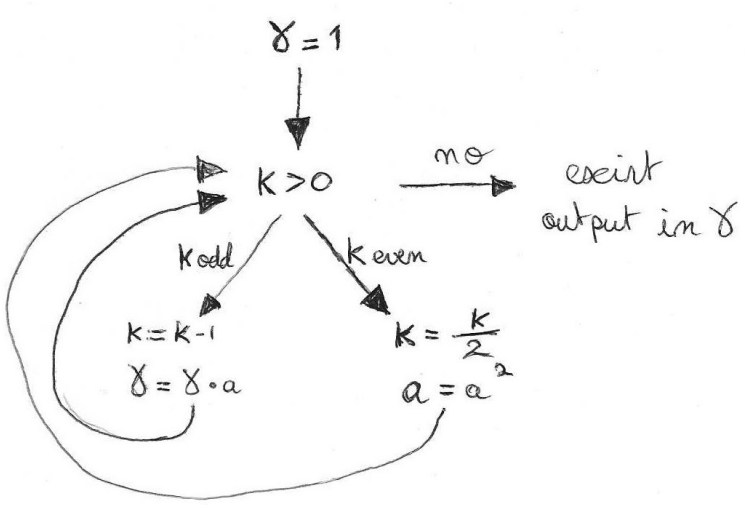

Figure 9.1: operation diagram

This gives $O(\log k)$ operations.

Whether decoding is simple or not depends on the realization of the cyclic group.

1. The group $(\mathbb{Z} \backslash n\mathbb{Z}, +)$ with generator $a = 1$. In this case the discrete logarithm problem is totally trivial. Given "a^l" means that we have l directly given. The group operations are residue class additions.

2. The multiplicative group $\mathbb{F}_q^*$ is also a cyclic group of order $(q-1)$ where $q = p^n$, $p \in \mathbb{P}$. It is more secure but q must be quite large. The group operations are polynomial multiplication combined with residue class multiplication.

3. Much better is the group $E(\mathbb{F}_q)$ of $\mathbb{F}_q$ valued points on an elliptic curve (suitable chosen). In general it will not be a cyclic group, but we take cyclic subgroups of it. The group operations are still reasonable simple but the discrete logarithm problem is hard. It is a very effective method. See Section 7.3 for an example.

9.2 Noetherian rings, the Hilbert's basis theorem.

Let R be a commutative ring with unit.

Definition 9.1. R is called *Noetherian* if every ideal $I \subset R$ is finitely generated, i.e. there exists $f_1, .., f_r \in R$ such that

$$I = (f_1, .., f_r) = \{\sum_{i=1}^{r} g_i \cdot f_i \mid g_i \in R\}\,.$$

Claim: The following statements are equivalent:

a) R is Noetherian.

b) Every ascending chain of ideals $I_1 \subset I_2 \subset I_3 \subset ...$ becomes stationary, i.e. there exists an n_0 such that $I_{n_0} = I_{n_0+1} = I_{n_0+2} = ...$

c) Every non-empty set of ideals in R has maximal elements.

Proof. **a) ⇒ b)** Let R be Noetherian and consider an arbitrary ascending chain of ideals $I_1 \subset I_2 \subset I_3 \subset ....$ Set $I = \bigcup_{i=1}^{\infty} I_i$ and notice that I is an ideal as well. Then $I = (f_1, .., f_r)$ for some $f_1, .., f_r \in R$. Each f_i belongs to some I_{n_i}. For $n = \max_{i=1,..,r} n_i$ we get $f_i \in I_n \quad \forall i = 1, .., r$. Therefore, $I_i \subset I \subset I_n, \ \forall i$, in particular $I_n = I_{n+1} = I_{n+2} = ...$.

b) ⇒ c) Assume that there exists a non-empty set M of ideals without a maximal element. Then for every $I_1 \in M$ there exists $I_2 \in M$ with $I_1 \subsetneqq I_2$. This gives an ascending chain of ideals $I_1 \subsetneqq I_2 \subsetneqq I_3 \subsetneqq ...$ that does not become stationary. This is a contradiction to b).

c) ⇒ a) Take an arbitrary ideal $I \subseteq R$. Consider the set M of all finitely-generated ideals with generators from I. By assumption M has maximal elements. We take one and call it I_0. As I_0 lies in M it is finitely generated, e.g. $I_0 = (f_1, .., f_{r_0})$ with $f_i \in I$. For every $f \in I$ we consider the ideal $J =$

$(f_1, .., f_{r_0}, f) \in M$. Now $I_0 \subset J$, and as I_0 is maximal under the finitely generated ones we obtain $I_0 = J$. Hence, $f \in I_0$. As this is true for all $f \in I$ we obtain $I = I_0$. In particular I, itself is finitely generated.

□

Recall that a *principle ideal domain (PID)* is a ring without zero divisors (a domain) for which all ideals can be generated by one element. Such ideals are called *principal ideals*. In particular, a PID is always Noetherian. Examples of PID, and hence of Noetherian rings, are given by: the integers $\mathbb{Z}$, the polynomial ring in one variable and the ring of formal power series $\mathbb{K}[[X]]$ in one variable. In fact, for $\mathbb{K}[[X]]$ the only ideals are the zero ideal and the ideals (X^n) with $n \in \mathbb{Z}_{\geqslant 0}$.

Starting from these rings the following theorem gives more examples of Noetherian rings.

Theorem 9.2. *(Hilbert's basis theorem) Let R be a Noetherian ring. Let $R[X]$ be the polynomial ring in one variable with coefficients from R. Then $R[X]$ is Noetherian.*

Proof. Suppose $R[X]$ is not Noetherian. Hence there exists an ideal I of $\mathbb{R}[X]$ that is not finitely generated. Let $f_1 \in I$ be a polynomial with smallest degree. We write $f_1 = a_1 X^{n_1} + ..., \text{with} \quad n_1 = \deg f_1$. Here ... means terms of degree lower than the one given explicitely. Since I is not finitely generated, there exists $f_2 \in I \backslash (f_1)$. Choose f_2 of smallest degree n_2 with: $f_2 = a_2 X^{n_2} + ...$ For chosen $f_1, ..., f_k$ we take $f_{k+1} \in I \backslash (f_1, ..., f_k)$ of smallest degree n_{k+1} with $f_{k+1} = a_{k+1} X^{k+1} + ...$ Clearly $n_1 \leq n_2 \leq ...$ by our construction.

Next we consider the chain of ideals in R generated by the coefficients $(a_1) \subset (a_1, a_2) \subset (a_1, a_2, a_3) \subset ....$ As R was assumed to be Noetherian, the chain becomes stationary, i.e. $\exists k : (a_1, ..., a_k) = (a_1, ..., a_k, a_{k+1})$. Hence, $a_{k+1} \in (a_1, ..., a_k)$ which says that $\exists b_1, .., b_k \in R$ such that $a_{k+1} = \sum_1^k b_i \cdot a_i$.

We set $g := f_{k+1} - \sum_{i=1}^k b_i \cdot f_i \cdot X^{n_{k+1} - n_i}$. Since $f_{k+1} \notin (f_1, ..., f_k)$ the same is true for g. However $\deg g < n_{k+1}$, which contradicts to our choice of f_{k+1}. Therefore, every ideal in $R[X]$ is finitely generated and $R[X]$ is thus Noetherian. □

Corollary 9.3. *If R is Noetherian then $R[X_1, ..., X_n]$ is Noetherian too.*

Proof. Notice $R[X_1, ..., X_n] \cong R[X_1, ..., X_{n-1}][X_n]$ and apply induction. □

By this we directly get

Theorem 9.4. *$\mathbb{Z}[X_1, ..., X_n]$ is Noetherian. $\mathbb{K}[X_1, ..., X_n]$ is Noetherian. In other words, the polynomial ring in n variables over a field $\mathbb{K}$ or over the integers $\mathbb{Z}$ is Noetherian.*

In particular, all ideals in the polynomial ring $\mathbb{K}[\underline{X}]$ are finitely generated.

This statement has some geometric consequences.

Theorem 9.5. *Let $\mathbb{K} \subset \mathbb{L}$ be a field extension. Then every descending chain $V_1 \supset V_2 \supset ..$ of affine varieties in $\mathbb{A}^n(\mathbb{L})$ defined over $\mathbb{K}$ in $\mathbb{A}^n(\mathbb{L})$ becomes stationary.*

Proof. Recall that to every variety V we can assign its vanishing ideal $\mathcal{I}(V)$, see Section 3.5.1. The descending chain of varieties give an ascending chain of ideals

$$\mathcal{I}(V_1) \subset \mathcal{I}(V_2) \subset ... \subset \mathbb{K}[X_1, ..., X_n] .$$

As the polynomial ring is Noetherian this chain becomes stationary, hence for some n, $\mathcal{I}(V_n) = \mathcal{I}(V_{n+1}) = ...$ But this implies

$$V_n = \mathfrak{V}(\mathcal{I}(V_n)) = \mathfrak{V}(I(V_{n=1}) = V_{n+1} ... .$$

□

In particular each affine variety has only finitely many affine subvarieties.

Zariski topology:

This is a topology defined for $\mathbb{A}^n(\mathbb{K}) = \mathbb{K}^n$. A topology on a set M can be given by defining which sets are the closed sets and taking care that certain basic axioms are fulfilled. These axioms are

1. M and $\emptyset$ are closed,
2. finite unions of closed sets are closed,
3. arbitrary intersections of closed sets are closed.

It is also possible to define a topology by defining what are the open set. Both definitions are related. The open sets will be the complements of the closed sets. In the axioms the role of union and intersection will be inverted.

In our case we declare the affine varieties $\mathfrak{V}(I) \subset \mathbb{A}^n(\mathbb{K}) = \mathbb{K}^n$ to be closed sets (hence their complements are declared open). Then the axioms of topology are satisfied as

- $\emptyset = \mathfrak{V}(1)$, $\mathbb{K}^n = \mathfrak{V}(0)$ are closed.

- $\bigcup_{i=1}^{r} \mathfrak{V}(I_i) = \mathfrak{V}(I_1 \cdot \cdots \cdot I_r)$ is closed.

- $\bigcap_i \mathfrak{V}(I_i) = \mathfrak{V}(\sum I_i)$ (for any number of $\mathfrak{V}(I_i)$) is closed.

With this definition $\mathbb{K}^n = \mathbb{A}^n(\mathbb{K})$ becomes a topological space. Note also that we can either take the definitions using vanishing sets of ideals or vanishing sets of finitely many polynomials. Due to the fact that the polynomial ring is Noetherian both definitions coincide.

Example. (Closed subsets in $\mathbb{A}^1(\mathbb{K}) = \mathbb{K}$). The relevant ring is the polynomial ring in one variable. Clearly $\mathbb{K} = \mathfrak{V}(0)$ and $\emptyset = \mathfrak{V}(1)$. Since all ideals in $\mathbb{K}[X]$ are principal, the only other closed sets are $\mathfrak{V}(f)$, $f \in \mathbb{K}[X]$ with f a non-constant polynomial in one variable. Since a polynomial of degree d can have at most d zeros, all such sets are finite. On the other hand, for every finite set of points $\{a_1, ..., a_r\} \in \mathbb{K}$, the set $Y = V((f))$ with

$$f(X) = (X - a_1)(X - a_2) \cdots (X - a_r)$$

consists exactly of these points, i.e. is an affine variety. Hence we obtain

Proposition 9.6. *The only Zariski closed sets in $\mathbb{K}$ are the sets $\emptyset$, $\mathbb{K}$, and its finite subsets.*

Bibliography

[Wer] Anette Werner, *Elliptische Kurven in der Kryptographie*, Springer 2002

[Kunz] Ernst Kunz, *Introduction to plane algebraic curves*, Birkhäuser 2005.

[KK] Ernst Kunz, *Commutatative Algebra*,

[RS] Martin Schlichenmaier, *An Introduction to Riemann Surfaces, Algebraic Surfaces and Moduli Spaces*, 2nd enlarged edition, Springer 2007

[Enge] Andreas Enge *Elliptic curves and their applications to cryptography* . Kluwer Academic Publishers 1999

[Hu] Klaus Hulek. *Elementare Algebraische Geometrie*, vieweg studium, 2000.

[HC] A. Hurwitz, R. Courant: *Allgemeine Funktionentheorie und elliptische Funktionen*, Springer, 1964.

[Bit] Ricardo Pérez-Marco: *Bitcoins and decentralised trust protocols*, Newsletter of the EMS, Vol. 100, 31–38 (2016)